

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :

**C12N 15/12, C07K 16/28, 19/00, C12N
5/10, 15/85, A61K 39/395****A1**

(11) International Publication Number:

WO 95/27061

(43) International Publication Date:

12 October 1995 (12.10.95)(21) International Application Number: **PCT/US95/04228**(22) International Filing Date: **4 April 1995 (04.04.95)**

(30) Priority Data:

08/222,616**4 April 1994 (04.04.94)****US**

(60) Parent Application or Grant

(63) Related by Continuation

US

Filed on

08/222,616 (CIP)**4 April 1994 (04.04.94)**(71) Applicant (for all designated States except US): **GENENTECH,
INC. [US/US]; 460 Point San Bruno Boulevard, South San
Francisco, CA 94080-4990 (US).**

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BENNETT, Brian, D.
[US/US]; 1332 Oddstad Boulevard, Pacifica, CA 94044
(US). GOEDEL, David [US/US]; 2115 Forestview Av-
enue, Hillsborough, CA 94010 (US). LEE, James, M.
[US/US]; 705 Shelter Creek, San Bruno, CA 94066 (US).
MATTHEWS, William [GB/US]; 560 Summit Springs
Road, Woodside, CA 94062 (US). TSAI, Siao, Ping [-/US];
519 Orange Avenue, South San Francisco, CA 94080 (US).****WOOD, William, I. [US/US]; 1400 Tarrytown Street, San
Mateo, CA 94402 (US).**(74) Agents: **LEE, Wendy, M. et al.; Genentech, Inc., 460 Point
San Bruno Boulevard, South San Francisco, CA 94080-4990
(US).**(81) Designated States: **CA, JP, MX, US, European patent (AT,
BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL,
PT, SE).****Published***With international search report.**Before the expiration of the time limit for amending the
claims and to be republished in the event of the receipt of
amendments.*(54) Title: **PROTEIN TYROSINE KINASE AGONIST ANTIBODIES**

(57) Abstract

Agonist antibodies are disclosed which bind to the extracellular domain of receptor protein tyrosine kinases pTKs, and thereby cause dimerization and activation of the intracellular tyrosine kinase domain thereof. The antibodies are useful for activating their respective receptor and thereby enabling the role of the tyrosine kinase receptor in cell growth and/or differentiation to be studied. Chimeric proteins comprising the extracellular domain of the receptor pTKs and an immunoglobulin constant domain sequence are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

BACKGROUND OF THE INVENTIONFIELD OF THE INVENTION

The present invention relates to novel protein tyrosine kinase (pTK)
5 genes, the proteins encoded by these genes, RNA nucleic acid sequences
which hybridize to the genes, antibodies specific for the encoded proteins,
chimeras of the proteins and methods of use therefor.

In particular, this application relates to agonist antibodies which
are able to activate the tyrosine kinase domain of the receptor pTKs.
10 disclosed herein and pTK-immunoglobulin chimeras.

DESCRIPTION OF RELATED ART

Transduction of signals that regulate cell growth and differentiation
is regulated in part by phosphorylation of various cellular proteins.
Protein tyrosine kinases are enzymes that catalyze this process. Moreover,
15 many act as growth factor receptors. The c-kit subgroup of receptor
tyrosine kinases catalyze the phosphorylation of exogenous substrates, as
well as tyrosine residues within their own polypeptide chains (Ullrich et
al., Cell 61:203 [1990]). Members of the c-kit subgroup include FLT/FLK
(Fetal Liver Kinase), FGF (Fibroblast Growth Factor Receptor) and NGF
20 (Nerve Growth Factor Receptor).

The EPH tyrosine kinase subfamily, Eph, Elk, Eck, Eek, Hek, Hek2,
Sek, Ehk-1, Ehk-2, Cek-4 to -10, Tyro 1, 4, 5 and 6, appears to be the
largest subfamily of transmembrane tyrosine kinases (Hirai et al., Science
238:1717-1720 [1987]; Letwin et al., Oncogene 3:621-678 [1988]; Lhotak et
25 al., Mol. Cell. Biol. 13:7071-7079 [1993]; Lindberg et al., Mol. Cell.
Biol. 10:6316-6324 [1990]; Bohme et al., Oncogene 8:2857-2862 [1993]; and
Wicks et al., Proc. Natl. Acad. Sci. USA. 89:1611-1615 [1992]; Pasquale et
al. Cell Regulation 2:523-534 [1991]; Sajjadi et al., New Biol. 3:769-778
[1991]; Wicks et al., Proc. Natl. Acad. Sci. USA. 89:1611-1615 [1992];
30 Lhotak et al., Mol. Cell. Bio. 11:2496-2502 [1991]; Gilardi-Hebenstreit et
al., Oncogene 7:2499-2506 [1992]; Lai et al., Neuron 6:691-704 [1991];
Sajjadi et al., Oncogene 8:1807-1813 [1993]; and Maisonpierre et al.,
Oncogene 8:3277-3288 [1993]).

Additional pTKs and agonist antibodies thereto are needed in order
35 to further study growth and differentiation of cells, for use as
therapeutic agents and for diagnostic purposes. Accordingly, it is an

object of the present invention to provide novel pTK genes, the proteins encoded thereby, antibodies specific for the encoded proteins, chimeras of the proteins and methods of use thereof.

SUMMARY OF THE INVENTION

5 The genes isolated as described herein are referred to, collectively, as "protein tyrosine kinase genes" or "pTK genes". The nucleic acid sequences of some of these genes, isolated as discussed herein, show significant homology with previously identified protein tyrosine kinases containing extracellular domains, which function as growth factor receptors
10 (e.g., pTKs of the c-kit subgroup). Some of the pTK genes have been shown to be present in both megakaryocytic and lymphocytic cells.

 In particular, fourteen pTK genes have been identified. Two pTK genes, referred to as SAL-S1 and SAL-D4 were identified in megakaryocytic cells. SAL-D4 is related to the CSK family of intracellular pTKs and SAL-S1
15 is related to the FGF receptor family of pTKs. Five pTK genes, referred to as LpTKs, were identified in lymphocytic cells and have been shown to be present in megakaryocytes as well. One pTK gene, referred to as HpTK5, was identified in human hepatoma cells. Six pTK genes, referred to as bpTK genes, were found in human brain tissue.

20 The pTK genes, which are the subject of the present invention, were generally identified using two sets of degenerative oligonucleotide primers: a first set which amplifies all pTK DNA segments (SEQ ID NOS: 1-2), and a second set which amplifies highly conserved sequences present in the catalytic domain of the c-kit subgroup of pTKs (SEQ ID NOS: 3-4). The
25 pTK genes identified in this manner are described below.

 SAL-S1 is expressed in several megakaryocytic cell lines, but not in erythroid cell lines. The nucleotide sequence of part of SAL-S1 was obtained, revealing a sequence containing 160 base pairs (SEQ ID NO: 5). This isolated DNA fragment encoded an amino acid sequence (SEQ ID NO: 6)
30 which exhibited significant sequence homology with known protein tyrosine kinases of the FLT/FLK family. The deduced amino acid sequence of SAL-S1 (SEQ ID NO: 32) contains 1298 residues.

 SAL-D4, also expressed in megakaryocytic cells, is a DNA fragment containing the nucleotide sequence of 147 base pairs. (SEQ ID NO: 7). This
35 isolated DNA fragment encoded an amino acid sequence (SEQ ID NO: 8) which exhibited significant sequence homology with known protein tyrosine kinases of the CSK intracellular pTK family.

The LpTKs, including LpTK 2, LpTK 3, LpTK 4, LpTK 13 and LpTK 25, are expressed in lymphocytic cells, as well as megakaryocytic cells. The nucleotide sequence (151 base pairs) of the LpTK 3 gene was obtained (SEQ ID NO: 11). The nucleotide sequences of the LpTK 2, LpTK 4, and LpTK 13 genes contained 149 base pairs (SEQ ID NO: 9), 137 base pairs (SEQ ID NO: 13), and 211 base pairs (SEQ ID NO: 15) respectively. LpTK 25 has a nucleotide sequence of 3120 b.p. (SEQ ID NO: 22). A full length gene sequence has been obtained for LpTK 2 (SEQ ID NO: 19) which contains 7607 b.p. Additional sequencing of LpTK 4 revealed a sequence of 404 b.p. (SEQ ID NO: 21).

The HpTK5 gene, expressed in human hepatoma cells, has a nucleotide sequence of 3969 b.p. (SEQ ID NO: 23).

Nucleotide sequences of the bpTKs, including bpTK 1, bpTK 2, bpTK 3, bpTK 4, bpTK 5 and bpTK 7, are expressed in human brain tissue and encode proteins having the amino acid sequences of SEQ ID NOS: 25-29 and 34 respectively.

Thus, the present invention includes DNA isolated from a human megakaryocytic cell line, which hybridizes to DNA encoding an amino acid sequence which is highly conserved in the catalytic domain of protein tyrosine kinases of the c-kit subgroup.

The present invention also includes the proteins encoded by the pTK genes identified as described herein, which exhibit significant sequence homology with members of the c-kit subgroup of pTKs as well as the proteins encoded by HpTK5 and the bpTKs. The present invention also includes SAL-S1, SAL-D4, LpTK, HpTK5 and bpTK homologues or equivalents (i.e., proteins which have amino acid sequences substantially similar, but not identical, to that of SAL-S1, SAL-D4, the LpTKs, HpTK5 and the bpTKs, which exhibit tyrosine kinase activity). This invention further includes peptides (SAL-S1, SAL-D4, LpTK, HpTK5 and bpTK fragments) which retain tyrosine kinase activity, yet are less than the entire SAL-S1, SAL-D4, LpTK, HpTK5 and bpTK sequences; and uses for the SAL-S1, SAL-D4, the LpTK, HpTK and the bpTK nucleic acid sequences and SAL-S1, SAL-D4, LpTK, HpTK and bpTK equivalents.

The present invention further includes nucleic acid sequences which hybridize with DNA or RNA encoding the proteins described herein, which exhibit significant sequence homology with the FLT/FLK, FGF receptor or NGF receptor family of protein tyrosine kinases contained within the c-kit subgroup. Such nucleic acid sequences are useful as probes to identify pTK genes in other vertebrates, particularly mammals, and in other cell types.

They can also be used as anti-sense oligonucleotides to inhibit protein tyrosine kinase activity, both *in vitro* and *in vivo*.

The SAL-S1, SAL-D4, LpTK, HpTK and bpTK tyrosine kinases of the present invention can be used as target proteins in conjunction with the development of drugs and therapeutics to modulate cell growth, differentiation and other metabolic functions. The SAL-S1, SAL-D4, LpTK, HpTK or bpTK proteins can be used as agonists or antagonists to other tyrosine kinases. The pTKs can also be instrumental in the modulation of megakaryocyte and/or platelet adhesion interactions.

10 In addition, the SAL-S1, SAL-D4, LpTK, HpTK and bpTK tyrosine kinases can be used in screening assays to detect cellular growth and/or differentiation factors. Using standard laboratory techniques, the ligands of the pTKs of the present invention can be identified. In particular, the invention provides chimeric pTK-immunoglobulin fusion proteins which are
15 useful for isolating ligands to the pTKs disclosed herein. The chimeric proteins are also useful for diagnostic assays designed to detect these ligands present endogenously, within cells, as well as exogenously, in extra-cellular fluids. Assays, using the chimeric proteins, can also be designed as diagnostic aids to detect these ligands in body fluids such as
20 blood and urine.

In another aspect, the invention provides antibodies specific for SAL-S1, SAL-D4, the LpTKs, HpTK5 and the bpTKs, which are optionally agonists for their respective pTK (where the pTK is a receptor). The invention also concerns a hybridoma cell line and an isolated nucleic acid
25 encoding a monoclonal antibody as herein defined.

Also, the invention pertains to a method for activating a pTK as herein disclosed, comprising reacting the pTK with an agonist antibody thereto. In a different aspect, the invention concerns a method for enhancing cell growth and/or differentiation comprising administering to
30 a human patient in need of such treatment a physiologically effective amount of an agonist antibody which activates a pTK as herein disclosed.

In a still further aspect, the invention concerns a method for detecting a pTK by contacting a source suspected of containing the pTK with a detectably labeled monoclonal antibody which reacts immunologically with
35 the pTK, and determining whether the antibody binds to the source.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B depict the nucleotide sequence of SAL-S1 (SEQ ID NO: 5) and its deduced amino acid sequence (SEQ ID NO: 6).

Figures 2A and 2B depict the nucleotide sequence of SAL-D4 (SEQ ID NO: 7) and its deduced amino acid sequence (SEQ ID NO: 8).

Figure 3A depicts the nucleotide sequence of LpTK 2 (SEQ ID NO: 9) and its deduced amino acid sequence (SEQ ID NO: 10).

Figure 3B depicts the nucleotide sequence of LpTK 3 (SEQ ID NO: 11) and its deduced amino acid sequence (SEQ ID NO: 12).

Figure 3C depicts the nucleotide sequence of LpTK 4 (SEQ ID NO: 13) and its deduced amino acid sequence (SEQ ID NO: 14).

Figure 3D depicts the nucleotide sequence of LpTK 13 (SEQ ID NO: 15) and its deduced amino acid sequence (SEQ ID NO: 16).

Figures 4A-4I depict the nucleotide sequence (SEQ ID NO: 17) of SAL-S1 and its deduced amino acid sequence (SEQ ID NO: 18).

Figures 5A-5K depict the full length nucleotide sequence (SEQ ID NO: 19) of LpTK2 and its deduced amino acid sequence (SEQ ID NO: 20).

Figure 6 depicts the partial nucleotide sequence (SEQ ID NO: 21) for LpTK4.

Figures 7A-7C depict the full length nucleotide sequence (SEQ ID NO: 22) for LpTK25.

Figures 8A-8I depict the full length nucleotide sequence (SEQ ID NO: 23) and the deduced amino acid sequence of HpTK5 (SEQ ID NO: 24).

Figure 9 depicts the amino acid sequence (SEQ ID NO: 25) of bpTK1.

Figure 10 depicts the amino acid sequence (SEQ ID NO: 26) of bpTK2.

Figure 11 depicts the amino acid sequence (SEQ ID NO: 27) of bpTK3.

Figure 12 depicts the amino acid sequence (SEQ ID NO: 28) of bpTK4.

Figure 13 depicts the amino acid sequence (SEQ ID NO: 29) of bpTK5.

Figure 14 depicts the amino acid sequence (SEQ ID NO: 30) of bpTK7.

Figures 15A-15F depict the full-length nucleotide sequence of SAL-S1 (SEQ ID NO: 31) and its deduced amino acid sequence (SEQ ID NO: 32).

Figures 16A-16H depict the full-length nucleotide sequence of bpTK7 (SEQ ID NO: 33) and its deduced amino acid sequence (SEQ ID NO: 34).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Novel protein tyrosine kinase genes have been identified, their nucleic acid sequences determined, and the amino acid sequences of the encoded proteins deduced. The genes isolated as described herein are

referred to, collectively, as "protein tyrosine kinase genes" or "pTK genes".

To facilitate the isolation and identification of these novel pTKs, two sets of DNA probes were used, as described in Example 1. The first set
5 generally consisted of two degenerative oligonucleotide sequences, pTK 1 (SEQ ID NO: 1) and pTK 2 (SEQ ID NO: 2) (Matthews, Cell 65:1143 [1991]; and Wilks, Proc. Natl. Acad. Sci. USA 86:1603 [1989]). These sequences were used as primers in a polymerase chain reaction to amplify tyrosine kinase DNA segments (Mullis, et al., Cold Spring Harbor Symp. Advan. Biol. 51:263
10 [1986]).

The second set generally consisted of two oligonucleotide sequences, pTK 3 (SEQ ID NO: 3) and pTKKW (SEQ ID NO: 4) designed to amplify the nucleic acid sequence which encodes the highly conserved regions of the catalytic domains of the c-kit family of protein tyrosine kinases. These
15 sequences were used as primers in the polymerase chain reaction (PCR) in a second round of DNA amplification. Using this two-step amplification procedure, DNA fragments which hybridized to these pTK primers were identified, isolated and subsequently sequenced.

In particular, fourteen pTK genes have been identified. Two pTK
20 genes, referred to as SAL-S1 and SAL-D4, were identified in several megakaryocytic cell lines, including CMK 11-5, DAMI, UT-7 and UT-7 grown in erythropoietin, but not in the erythroid cell lines HEL, PMA stimulated HEL cells, or K562. Five pTK genes, referred to as LpTKs, were identified in lymphocytic, as well as in megakaryocytic cells. One pTK gene, referred
25 to as HpTK5, was identified in human hepatoma cells, and six genes, referred to as bpTKs, were identified in human brain tissue.

SAL-S1 (SEQ ID NOS: 6, 18 and 32) encoded by the nucleic acid sequence of SEQ ID NOS: 5, 17 and 31 exhibits significant homology with the FLT/FLK family of pTKs. SAL-S1 has a signal peptide (i.e., amino acid
30 residues 1 to 24 of Figure 15); extracellular domain (i.e., amino acid residues 25 to 775 of Figure 15); transmembrane domain (i.e., amino acid residues 776 to 800 of Figure 15) and a cytoplasmic tyrosine kinase domain (i.e., amino acid residues 801 to 1298 of Figure 15). SAL-D4 (SEQ ID NO: 8) encoded by SEQ ID NO: 7 is related to the CSK family of intracellular
35 pTKs. The LpTKs, LpTK 2 (SEQ ID NOS: 10 and 20) encoded by SEQ ID NOS: 9 and 19; LpTK 3 (SEQ ID NO: 12) encoded by SEQ ID NO: 11; LpTK4 (SEQ ID NO: 14) encoded by SEQ ID NOS: 13 and 21; LpTK13 (SEQ ID NO: 16) encoded by SEQ

ID NO: 15; and LpTK25 encoded by SEQ ID NO: 22, also exhibit sequence homology with known protein tyrosine kinases.

HpTK5 (SEQ ID NO: 24) encoded by SEQ ID NO: 23 and the bpTKs 1, 2, 3, 4, 5 and 7 (SEQ ID NOS: 25-29 and 34 respectively), similarly exhibit sequence homology with known protein tyrosine kinases. BpTK7 encodes a receptor pTK with a signal peptide (i.e., amino acid residues 1-19 of Figure 16); extracellular domain (i.e., amino acid residues 20-547 of Figure 16); and transmembrane domain (i.e., amino acid residues 548-570 of Figure 16). The remaining sequence comprises the intracellular tyrosine kinase domain.

Thus, as described above, DNA molecules which hybridize with DNA encoding amino acid sequences present in the catalytic domain of a protein tyrosine kinase of the c-kit subgroup of protein kinases have been isolated and sequenced. These isolated DNA sequences, collectively referred to as "pTK genes", (and their deduced amino acid sequences) have been shown to exhibit significant sequence homology with known members of pTK families.

Once isolated, these DNA fragments can be amplified using known standard techniques such as PCR. These amplified fragments can then be cloned into appropriate cloning vectors and their DNA sequences determined.

These DNA sequences can be excised from the cloning vectors, labeled with a radiolabeled nucleotide such as ³²P and used to screen appropriate cDNA libraries to obtain the full-length cDNA clone.

The pTK genes as described above have been isolated from the source in which they occur naturally, e.g., megakaryocytic and lymphocytic cells. The present invention is intended to include pTK genes produced using genetic engineering techniques, such as recombinant technology, as well as pTK genes that are synthesized chemically.

The deduced amino acid sequences of the pTK genes include amino acid sequences which encode peptides exhibiting significant homology with the catalytic domain of protein tyrosine kinases of the c-kit subgroup of tyrosine kinases. These proteins, encoded by the pTK genes, can include sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence, resulting in a silent change, that is a change not detected phenotypically. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent substitution.

In addition, the protein structure can be modified by deletions, additions, inversion, insertions or substitutions of one or more amino acid residues in the sequence which do not substantially detract from the desired functional tyrosine kinase properties of the peptide.

5 Modified pTKs of the present invention, with tyrosine kinase activity, can be made using recombinant DNA techniques, such as excising it from a vector containing a cDNA encoding such a protein, or by synthesizing DNA encoding the desired protein mechanically and/or chemically using known techniques.

10 An alternate approach to producing the pTKs of the present invention is to use peptide synthesis to make a peptide or polypeptide having the amino acid sequence of such a protein, depending on the length of the pTK desired. The peptides or modified equivalents thereof, can be synthesized directly by standard solid or liquid phase chemistries for peptide
15 synthesis.

Preferably, the pTKs of the present invention will be produced by inserting DNA encoding the proteins into an appropriate vector/host system where it will be expressed. The DNA sequences can be obtained from sources in which they occur naturally, can be chemically synthesized or can be
20 produced using standard recombinant technology.

This invention also pertains to an expression vector comprising a pTK gene of the present invention, encoding for a protein which exhibits receptor tyrosine kinase activity.

The pTK genes of the present invention can be used for a number of
25 diagnostic and therapeutic purposes. For example, the nucleic acid sequences of the pTK genes can be used as probes to identify other protein tyrosine kinases present in other cell types, including eukaryotic and prokaryotic cell types.

The nucleic acid sequences can also be used to design drugs that
30 directly inhibit the kinase activity of protein tyrosine kinases, or to design peptides that bind to the catalytic domain of tyrosine kinases, thus inhibiting their activity. These sequences can also be used to design anti-sense nucleotides that can also inhibit, or destroy, tyrosine kinase activity. Such inhibition of tyrosine kinase activity would be desirable
35 in pathological states where decreased cellular proliferation would be beneficial, such as leukemias or other malignancies.

The nucleic acid sequences can also be used to design drugs, peptides or anti-sense nucleotides as above, but with enhancing, rather than

inhibitory effects, on tyrosine kinases. Such enhanced tyrosine kinase activity would result in increasing the phosphorylation of substrates (exogenous, as well as endogenous tyrosine residues). Enhanced effects would be desirable in states where increased cellular proliferation would be beneficial, such as anemias, bleeding disorders and during surgical procedures.

The pTK genes of the present invention can also be used to obtain soluble fragments of receptor tyrosine kinases, capable of binding their respective ligands. pTK genes encoding soluble tyrosine kinase fragments can be produced using recombinant DNA techniques or synthetically. In either case, the DNA obtained encodes a soluble pTK fragment which lacks a substantial portion of the hydrophobic transmembrane region to permit solubilization of the fragment.

These soluble pTK protein fragments can be introduced exogenously to act as competitors with the endogenous, membrane bound pTK for their respective ligands, thus inhibiting tyrosine kinase activity. Alternately, a modified soluble pTK protein fragment can be introduced which binds the ligand but does not activate kinase activity.

These soluble pTK protein fragments can also be used in binding assays to detect ligands such as growth and differentiation factors. Once these ligands are identified, they may be altered or modified to inhibit or enhance kinase activity. For example, the ligands may be modified or attached to substances that are toxic to the cell, such a ricin, thus destroying the target cell. The substance may be a super-activating substance which, after binding to the pTK, may substantially increase the kinase activity, or activate other growth factors.

pTK genes of the present invention would also be useful to develop diagnostic tools for in vitro screening assays for ligands such as growth factors or differentiation factors that inhibit or enhance kinase activity. The proteins encoded by the pTK genes can also be used in such assays, or as immunogens to produce monoclonal or polyclonal antibodies to be used in such assays.

In one embodiment of the invention, a chimera comprising a fusion of the extracellular domain of the pTK (where the pTK is a receptor) and an immunoglobulin constant domain can be constructed which can be used to assay for ligands for the receptor and can be used for the production of antibodies against the extracellular domain of the receptor.

The expression "extracellular domain" or "ECD" when used herein refers to any polypeptide sequence that shares a ligand binding function of the extracellular domain of the naturally occurring receptor pTKs disclosed herein. Ligand binding function of the extracellular domain refers to the ability of the polypeptide to bind at least one pTK ligand. Accordingly, it is not necessary to include the entire extracellular domain since smaller segments are commonly found to be adequate for ligand binding. The truncated extracellular domain is generally soluble. The term ECD encompasses polypeptide sequences in which the hydrophobic transmembrane sequence (and, optionally, 1-20 amino acids C-terminal and/or N-terminal to the transmembrane domain) of the mature pTK has been deleted. Thus, the soluble extracellular domain-containing polypeptide can comprise the extracellular domain and the cytoplasmic domain of the pTK. Alternatively, in the preferred embodiment, the polypeptide comprises only the extracellular domain of the pTK. The extracellular and transmembrane domains of the pTK can be readily determined by the skilled practitioner by aligning the pTK of interest with known pTK amino acid sequences for which these domains have been delineated. Alternatively, the hydrophobic transmembrane domain can be readily delineated based on a hydrophobicity plot of the sequence. The extracellular domain is N-terminal to the transmembrane domain.

The term "immunoglobulin" generally refers to polypeptides comprising a light or heavy chain usually both disulfide bonded in the native "Y" configuration, although other linkage between them, including tetramers or aggregates thereof, is within the scope hereof.

Immunoglobulins (Ig) and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Patent 4,745,055; EP 256,654; Faulkner et al., Nature 298:286 [1982]; EP 120,694; EP 125,023; Morrison, J. Immunol. 123:793 [1979]; Köhler et al., Proc. Nat'l. Acad. Sci. USA 77:2197 [1980]; Raso et al., Cancer Res. 41:2073 [1981]; Morrison et al., Ann. Rev. Immunol. 2:239 [1984]; Morrison, Science 229:1202 [1985]; Morrison et al., Proc. Nat'l. Acad. Sci. USA 81:6851 [1984]; EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. patent 4,444,878; WO 88/03565; and EP 68,763 and references cited therein. The immunoglobulin moiety in the chimera of the present invention may be obtained from IgG₁, IgG₂, IgG₃, or IgG₄ subtypes, IgA, IgE, IgD or IgM, but

preferably IgG₁ or IgG₃. Most preferably, the immunoglobulin moiety is the Fc portion of IgG-γ.

The terms "chimera comprising a fusion of an extracellular domain of a pTK with an immunoglobulin constant domain sequence" or "pTK-immunoglobulin chimera" refer to a polypeptide comprising an extracellular domain coding amino acid sequence of a pTK conjugated to an immunoglobulin constant domain sequence. This definition includes chimeras in monomeric, homo- or heteromultimeric, and particularly homo- or heterodimeric, or tetrameric forms.

A preferred embodiment is the fusion of the C-terminus of the extracellular domain of a pTK, to the N-terminus of the C-terminal portion of an antibody (in particular the Fc domain), containing the effector functions of immunoglobulin G₁. In a preferred embodiment, the entire heavy chain constant region is fused to the extracellular domain. In another preferred embodiment, a sequence beginning in the hinge region just upstream of the papain cleavage site (which defines IgG Fc chemically; residue 216, taking the first residue of heavy chain constant region to be 114 (Kabat et al., Sequences of Immunological Interest, National Institutes of Health, Bethesda, MD, [1987])), or analogous sites of other immunoglobulins) is fused to the ECD of the pTK.

In a particularly preferred embodiment, the pTK extracellular domain is fused to the hinge region and C_H2 and C_H3 or C_H1, hinge, C_H2 and C_H3 domains of an IgG₁, IgG₂ or IgG₃ heavy chain. The precise site at which the fusion is made is not critical, and the optimal site can be determined by routine experimentation. A principal advantage of the chimeras is that they are secreted into the culture medium of recombinant hosts, although the degree of secretion might be different for various expression systems.

In general, the chimeras of the present invention are constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; EP 173,494; Munro, Nature 312: [13 December 1984]; Neuberger et al., Nature 312: [13 December 1984]; Sharon et al., Nature 309: [24 May 1984]; Morrison et al., Proc. Nat'l. Acad. Sci. USA 81:6851-6855 [1984]; Morrison et al. Science 229:1202-1207 [1985]; Boulianne et al., Nature 312:643-646 [13 December 1984]; Capon et al., Nature 337, 525-531 [1989]; Traunecker et al., Nature 339, 68-70 [1989].

To prepare the pTK-Ig chimeric polypeptides, the DNA including a region encoding the desired pTK sequence is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the immunoglobulin-like domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature pTK (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the pTK (where the native signal is employed). This DNA fragment then is readily inserted proximal to DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, the resulting construct tailored by deletional mutagenesis.

5

10 Preferably, the Ig is a human immunoglobulin when the variant is intended for in vivo therapy for humans. DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See for example, Adams et al., Biochemistry 19:2711-2719 [1980]; Gough et al., Biochemistry 19:2702-2710 [1980]; Dolby et al., P.N.A.S. USA, 77:6027-6031 [1980]; Rice et al., P.N.A.S. USA 79:7862-7865 [1982]; Falkner et al., Nature 298:286-288 [1982]; and Morrison et al., Ann. Rev. Immunol. 2:239-256 [1984].

15

The chimeric proteins disclosed herein are useful as diagnostics for isolating or screening ligands for the pTK of interest using the techniques of Lyman et al., Cell 75:1157-1167 [1993], for example. Also, the chimeric proteins are useful for diagnostic purposes for studying the interaction of various ligands with the extracellular domain of the various pTKs (see, e.g., Bennett et al., J. Biol. Chem. 266(34):23060-23067 [1991]). The chimeric proteins are further useful for the production of antibodies against the extracellular domain of the pTK (see Examples 3 and 5 herein). The chimeric proteins also have an additional therapeutic utility insofar as they provide a soluble form of the extracellular domain of the pTK which generally has an enhanced plasma half life (compared to the extracellular domain only) and therefore can be formulated in a pharmaceutically acceptable carrier and administered to a patient. The chimeric proteins are believed to find use as therapeutic agents for removal of excess systemic or tissue-localized pTK ligand which has been administered to a patient. Removal of excess ligand is particularly desirable where the ligand may be toxic to the patient. The chimeric protein acts to bind the ligand in competition with the endogenous pTK in the patient. Similarly, it is contemplated that the chimeric protein can be administered to a patient simultaneously, or subsequent to, administration of the ligand in the form of a sustained release composition. The chimeric protein acts as a soluble

20

25

30

35

binding protein for the ligand, thereby extending the half-life of the ligand.

The term "antibody" is used herein in the broadest sense and specifically covers polyclonal antibodies, monoclonal antibodies, immunoglobulin chains or fragments thereof, which react immunologically with a pTK.

In the preferred embodiment of the invention, the antibodies are monoclonal antibodies produced using techniques which are well known in the art. For example, the hybridoma technique described originally by Kohler and Milstein, Eur. J. Immunol., 6:511 [1976], and also described by Hammerling et al., In: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 [1981] can be used. The techniques of Cote et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies [Cote et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 [1985] and Boerner et al., J. Immunol., 147(1):86-95 [1991]].

The term "monoclonal antibody" as used herein refers to an antibody (as hereinabove defined) obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they can be synthesized by a hybridoma culture, uncontaminated by other immunoglobulins.

"Humanized" forms of non-human (e.g., murine) antibodies are immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂, or other antigen-binding subsequences of antibodies) which contain minimal amino acid residues derived from a non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework region (FR) residues of the human immunoglobulin are replaced

by corresponding non-human FR residues. Furthermore, a humanized antibody may comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. These modifications are made to further refine and optimize antibody performance.

5 The monoclonal antibodies herein include hybrid (chimeric) and recombinant antibodies produced by splicing a variable (including hypervariable) domain of an anti-pTK antibody with a constant domain (e.g., "humanized" antibodies), only one of which is directed against a pTK, or a light chain with a heavy chain, or a chain from one species with a chain
10 from another species, or fusions with heterologous proteins, regardless of species of origin or immunoglobulin class or subclass designation, so long as they are able to bind to the pTK of interest [See, e.g., Cabilly, et al., U.S. Patent No. 4,816,567; and Mage & Lamoyi, in Monoclonal Antibody Production Techniques and Applications, pp.79-97 (Marcel Dekker, Inc., New
15 York [1987])].

For "chimeric" and "humanized" antibodies see, for example, U.S. Patent No. 4,816,567; WO 91/09968; EP 452,508; and WO 91/16927.

Thus, the modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of
20 antibodies, and is not to be construed as requiring production of the antibody by any particular method.

In the most preferred embodiment of the invention, the antibodies are agonist antibodies. By "agonist antibody" is meant an antibody which is able to bind to, and activate, a particular pTK. For example, the agonist
25 may bind to the extracellular domain of the pTK and thereby cause dimerization of the pTK, resulting in transphosphorylation and activation of the intracellular catalytic kinase domain. Consequently, this may result in stimulation of growth and/or differentiation of cells expressing the receptor *in vitro* and/or *in vivo*. The agonist antibodies herein are
30 preferably against epitopes within the extracellular domain of the pTK, and preferably have the same biological characteristics as the monoclonal antibody produced by the hybridoma cell line deposited under American Type Culture Collection Accession No. ATCC HB 11,583. By "biological characteristics" is meant the *in vitro* and/or *in vivo* activities of the
35 monoclonal antibody, e.g., ability to activate the kinase domain of a particular pTK, ability to stimulate cell growth and/or differentiation of cells expressing the pTK, and binding characteristics of the antibody, etc. Accordingly, the antibody preferably binds to substantially the same

epitope as the anti-HpTK5 monoclonal antibody specifically disclosed herein. Most preferably, the antibody will also have substantially the same or greater antigen binding affinity of the anti-HpTK5 monoclonal antibody disclosed herein. To determine whether a monoclonal antibody has
5 the same specificity as the anti-HpTK5 antibody specifically disclosed (i.e., the antibody having the ATCC deposit No. HB 11,583), one can, for example, use a competitive ELISA binding assay.

DNA encoding the monoclonal antibodies useful in the method of the invention is readily isolated and sequenced using conventional procedures
10 (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as E. coli cells,
15 simian COS cells, Chinese Hamster Ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells.

The agonist antibodies disclosed herein are useful for *in vitro* diagnostic assays for activating the pTK receptor of interest. This is
20 useful in order to study the role of the receptor in cell growth and/or differentiation.

The pTK agonist antibodies have a further therapeutic utility in a method for enhancing cell growth and/or differentiation comprising administering to a human patient in need of such treatment a
25 physiologically effective amount of an exogenous pTK agonist antibody. Agonist antibodies to the SAL-S1 pTK may find utility in treating bleeding disorders and anemias, since this pTK was found to be expressed in megakaryocytic cells. The bpTK agonist antibodies may similarly be used to enhance differentiation and/or proliferation of brain cells in
30 neurodegenerative diseases (such as Alzheimers disease) based on the expression of these receptors in brain tissue. Finally, HpTK5 agonist antibodies may be used to enhance proliferation of primitive hematopoietic cells in patients having undergone chemo- or radiation therapy or bone marrow transplantation.

35 An "exogenous" therapeutic compound is defined herein to mean a therapeutic compound that is foreign to the mammalian patient, or homologous to a compound found in the mammalian patient but produced outside the mammalian patient.

The antibodies of the present invention are also suitable for detecting a pTK by contacting a source suspected to contain the pTK with a detectably labeled monoclonal antibody, and determining whether the antibody binds to the source. There are many different labels and methods of labeling known in the art. Suitable labels include, for example, enzymes, radioisotopes, fluorescent compounds, chemi- and bioluminescent compounds, paramagnetic isotopes. The pTK may be present in biological samples, such as biological fluids or tissues. For analytical or diagnostic purposes, the antibodies of the present invention are administered in an amount sufficient to enable the detection of a site on a pTK for which the monoclonal antibody is specific. The concentration of the detectably labeled monoclonal antibody should be sufficient to give a detectable signal above background, when bound to a pTK epitope.

The pTK agonist antibodies disclosed herein may be administered to a mammal, preferably a human, in a pharmaceutically acceptable dosage form, including those that may be administered to a human intravenously as a bolus or by continuous infusion over a period of time, by intramuscular, subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation routes.

Such dosage forms encompass pharmaceutically acceptable carriers that are inherently nontoxic and nontherapeutic. Examples of such carriers include ion exchangers, alumina, aluminum stearate, lecithin, serum proteins, such as human serum albumin, buffer substances such as phosphates, glycine, sorbic acid, potassium sorbate, partial glyceride mixtures of saturated vegetable fatty acids, water, salts, or electrolytes such as protamine sulfate, disodium hydrogen phosphate, potassium hydrogen phosphate, sodium chloride, zinc salts, colloidal silica, magnesium trisilicate, polyvinyl pyrrolidone, cellulose-based substances, and polyethylene glycol. Carriers for topical or gel-based forms of antibody include polysaccharides such as sodium carboxymethylcellulose or methylcellulose, polyvinylpyrrolidone, polyacrylates, polyoxyethylene-polyoxypropylene-block polymers, polyethylene glycol, and wood wax alcohols. For all administrations, conventional depot forms are suitably used. Such forms include, for example, microcapsules, nano-capsules, liposomes, plasters, inhalation forms, nose sprays, and sublingual tablets. The antibody will typically be formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml.

Pharmaceutical compositions may be prepared and formulated in dosage forms by methods known in the art; for example, see Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pennsylvania, 15th Edition 1975.

5 An effective amount of the pTK agonist antibody to be employed therapeutically will depend, for example, upon the therapeutic objectives, the route of administration, and the condition of the patient. Accordingly, it will be necessary for the therapist to titer the dosage and modify the route of administration as required to obtain the optimal
10 therapeutic effect. A typical daily dosage might range from about 1 μ g/kg to up to 1000 mg/kg or more, depending on the factors mentioned above. Typically, the clinician will administer the molecule until a dosage is reached that achieves the desired effect. The progress of this therapy is easily monitored by conventional assays.

15 Depending on the type and severity of the disease, from about 0.001 mg/kg to about 1000 mg/kg, more preferably about 0.01 mg to 100 mg/kg, more preferably about 0.010 to 20 mg/kg of the agonist antibody might be an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous
20 infusion. For repeated administrations over several days or longer, depending on the condition, the treatment is repeated until a desired suppression of disease symptoms occurs or the desired improvement in the patient's condition is achieved. However, other dosage regimens may also be useful.

25 The present invention will now be illustrated by the following Examples, which are not intended to be limiting in any way. The disclosures of all literature references cited in the specification are expressly incorporated herein by reference.

EXAMPLE 1

30 IDENTIFICATION AND ISOLATION OF pTK GENES

To facilitate the isolation and identification of these novel pTK genes, two sets of DNA probes were generally used (see Table 1).

The first set consisted of two degenerate oligonucleotide sequences, pTK 1 (SEQ ID NO: 1) and pTK 2 (SEQ ID NO: 2). These sequences were used
35 as polymerase chain reaction (PCR) primers, using standard PCR techniques, to amplify tyrosine kinase DNA segments.

10

First Round of Amplification

15

pTK3 (kit family specific) 5'-CGGATCCATCCACAGAGATGT-3'

pTKKW (kit family specific) 5'-GGAATTCCTTCAGGAGCCATCCACTT-3'

ISOLATION AND CHARACTERIZATION OF HpTK5

20

-18-

2.5 units Taq polymerase (Perkin-Elmer/Cetus) and 50pmol each of pTK-specific degenerate primers

[pTK1 5'TCGGATCCACA/CGNGAC/TC/TTGGC 3' (SEQ ID NO. 35),

pTK1B 5'TCGGATCCAC/TC/AGNGAC/TC/TTNGCNGC 3' (SEQ ID NO. 36),

5 pTK2 5'CTCGAATTCCA/GA/TAA/GC/GT/ACCAG/CACA/GTC 3' (SEQ ID NO. 37),

pTK2B 5'CTCGAATTCCA/GA/TAT/CC/GT/ACCAT/AACA/GTC 3'(SEQ ID NO. 38)]

derived from consensus regions among known pTKs as previously reported by others (Hanks et al., Science, 241:42-52 [1988]; Wilks, Proc. Nat. Acad. Sci., USA 86:1603-1607 [1989]; and Matthews et al., Cell 65:1143-

10 1152 [1991]). The PCR cycle was 1.5min at 95°C, 2min at 37°C and 3 min at 63°C repeated 35 times. The reaction product was electrophoretically separated on a 2% low-melting agarose gel, purified on an Elutip-D column (Schleicher & Schuell) digested with EcoR1 and BamH1, and subcloned into pUC19.

15 Recombinants were sequenced by the Sanger dideoxy method and evaluated by the FASTA nucleic acid sequence analysis program. One clone termed HpTK5 (214 bp) was radiolabelled by random priming and used to screen an oligo dT-primed lambda gt10 Hep3B cDNA library. DNA was isolated from 17 positive phage plaques and inserts were subcloned into
20 the EcoR1 site of pBluescript (Stratagene La Jolla, CA). The largest insert, a 3969 bp cDNA, was sonicated to an average size of 800-2000 bp and cloned into the Sma1 site of M13. Overlapping clones were sequenced using the Taq Dye Primer Cycle Method (CABI) on the Catalyst 800 Molecular Biology Lab Station (ABI). Sequencing reactions were then
25 analyzed on the ABI 373A Automated DNA Sequenator.

A single full-length 3969 bp cDNA was isolated and sequenced. (Figures 8A-8F). The full length clone, named hepatoma transmembrane kinase (HTK) or HpTK5, included an open reading frame extending from nucleotide 90 to 3050 predicted to encode a 987 amino acid protein of
30 108,270 Dalton. The putative initiation codon is preceded by an in-frame stop codon beginning at base 78. Preceding the open reading frame is a 5' untranslated region which is GC-rich as is characteristic for many growth factors or growth factor receptors (Kozak, J. Cell Biol. 115:887-903 [1991]).

35 The predicted protein sequence includes a transmembrane region (aa 538-563) which divides HpTK5 into extracellular (ECD) and intracellular domains (ICD). The ECD of 538 amino acids includes a signal peptide of 15 amino acids and a cysteine-rich box containing 20 Cys residues. In

addition, there are two fibronectin type III repeats spanning aa 321 to 425 and 435 to 526. Asn at positions 208, 340 and 431 are possible sites for N-glycosylation.

The putative intracellular domain (ICD) contains a kinase consensus region from position 613 through 881. This kinase region includes a putative ATP-binding consensus (Gly-X-Gly-X-X-Gly) in subdomain I at positions 622-627. A Lys at position 647 (subdomain II) corresponds to an invariant Lys among tyrosine kinases thought to be critical for the phosphotransfer reaction. Signature regions indicative of substrate specificity suggest that HpTK5 is a tyrosine rather than a serine/threonine kinase. These include the sequence at positions 740-745 in subdomain VI and the sequence at positions 783-790 in subdomain VIII. Tyrosine residues at positions 601, 619 and 741 are possible substrates for tyrosine kinase activity.

The predicted amino acid sequence of HpTK5 most closely resembles that of the subfamily originally defined by *EPH*. The pattern of expression of the *EPH* subfamily is suggestive of a role in differentiation and development. In particular, the emergence of neural elements corresponds with the expression of certain *EPH*-related genes. The *EPH* family receptors, Hek2 and Elk, are the most closely related pTKs to HpTK5. They share 79.3 and 76.5% identity within the ICD respectively and 45 and 42% identity within the ECD respectively.

B. Chromosome Mapping of HpTK5

Somatic cell hybrid DNAs from a panel of 25 human-hamster cell lines (Bios, New Haven, CN) were used for chromosome localization by PCR. Two sets of primers from the 3' untranslated region of HpTK5 were chosen. PCR was performed with 250 ng DNA and 50 pmol each of the 5' and 3' primers, 50 mM KCl, 1.5mM MgCl₂, 20 µg/ml gelatin, 0.2 mM dNTPs and 2.5 units Taq polymerase in a final volume of 100 µl. Cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec were repeated 30 times. A portion of each sample (15 µl) was electrophoresed through a 1.5% agarose gel, transferred to a nylon membrane and hybridized to a ³²P-labelled full length HpTK5 cDNA probe prior to 5 hour autoradiography. Positives were scored and compared to a matrix summary of human chromosomal material present in each of the somatic cell hybrid DNAs.

The 3'-untranslated region characteristically contains few, if any, intervening sequences and has a high degree of diversity among members

of gene families making it preferred in this type of analysis. Both sets of primers gave results that were consistent with human chromosome 7 only. Human chromosome 7 also includes the genes for the EGF receptor, hepatocyte growth factor (HGF) receptor, HGF, platelet-derived growth factor (PDGF) and interleukin-6. Karyotypic abnormalities involving this chromosome are common among human leukemias, particularly in aggressive myeloid leukemias that occur following radiation, alkylating agent chemotherapy or a pre-existing myelodysplastic condition (Baer et al., Curr. Opin. Oncol. 4:24-32 [1992]).

10 C. Northern Blotting of HpTK5

Poly-A selected RNA was electrophoresed through a 1.2% agarose, 2.2M formaldehyde gel and transferred to a nylon filter. Prepared or commercially obtained filters were hybridized in 50% formamide at 42°C to ³²-P labeled HpTK5, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or actin cDNA inserts and washed under stringent conditions (final wash: 0.1 x SSC, 0.2% SDS at 65°C). SSC is 0.15 M NaCl/ 0.015M Na₃citrate, pH 7.6. Northern blots of human fetal or adult tissue RNA were obtained from Clontech (Palo Alto, CA) and contained 2 µg/lane of poly A selected RNA.

Northern blot analysis of human fetal tissues revealed a single transcript of ~4Kb in heart, lung, liver and kidney, with a lesser signal detectable in brain. In adult human tissue, no signal was detectable in brain, while placenta had a particularly intense signal followed by kidney, liver, lung and pancreas. Skeletal muscle and heart were of lower signal intensity.

HpTK5 expression in human tumor cell lines was also analyzed by Northern blot analysis performed as discussed above. Cell lines derived from liver, breast (MCF 7), colon (Colo 205), lung (NCI 69), melanocyte (HM-1) or cervix (HeLa) had detectable signal of appropriate size. Message was present in select cell lines of hematopoietic origin. K562 (a primitive myeloid cell with multipotential), THP-1 (a monocytoid cell), U937 (a myelomonocytic cell line), Hep3B (a human hepatocarcinoma cell line), and CMK (of megakaryocytic origin) were all positive for HpTK5 message, but lymphoid (H9, Jurkat, JH-1, Raji, Ramos) or select other myeloid cells (KG-1 or KMT2) had no detectable transcript by Northern analysis.

Differential expression of the HpTK5 transcript in fetal versus adult brain suggests that HpTK5 may share, with other EPH subfamily

members, a role in events related to neural development. However, unlike some members of the *EPH* subfamily which are exclusively expressed in neurons (Maisonpierre et al., *supra*), HpTK5 is widely expressed in other tissues. In particular, HpTK5 is expressed in hematopoietic cells including CD34+ hematopoietic progenitor cells. The presence of the HpTK5 message in early hematopoietic cells and cell lines of myeloid lineage, but not in cell lines derived from lymphoid cells, suggests that HpTK5 may have lineage restricted expression.

EXAMPLE 3

10 PRODUCTION OF POLYCLONAL ANTIBODIES TO HPTK5

An HpTK5 extracellular domain (ECD)-human IgG₁ Fc fusion gene was constructed and fusion protein produced as previously described (Bennett et al., *J. Biol. Chem.* 266:23060-23067 [1991]). Polyclonal antibodies were generated in New Zealand White rabbits against the fusion protein; 4 µg in 100 µL PBS was emulsified with 100 µL Freund's adjuvant (complete adjuvant for the primary injection and incomplete adjuvant for all boosts). For the primary immunization and the first boost, the protein was injected directly into the popliteal lymph nodes (Sigel et al., *Methods Enzymol.* 93:3-12 [1983]). For subsequent boosts, the protein was injected into subcutaneous and intramuscular sites. 1.3 µg protein/kg body weight was injected every 3 weeks with bleeds taken 1 and 2 weeks following each boost. HpTK5 specificity of the immunized rabbit serum was assessed by flow cytometric analysis of NIH3T3 cells transfected with full length HpTK5 or vector alone using a 1:200 dilution of pre-immune serum or anti-HpTK5-IgG Fc serum. Significant peak shifts were observed in several HpTK5 expressing clones as compared to either pre-immune serum or vector alone transfectant controls.

EXAMPLE 4

UTILITY AND AGONIST ACTIVITY OF POLYCLONAL ANTIBODIES TO HPTK5

30 A. FLAG-HpTK5 Fusion Construct

Overlapping oligonucleotides encoding a 12 amino acid peptide having the sequence MDYKDDDDKKLAM (SEQ ID NO: 39) which includes the 4 amino acid antibody recognition site "FLAG" (IBI, New Haven, CT) a 5'-EcoRV restriction site and a 3'-NcoI restriction site

(5'-CCGGATATCATGGACTACAAGGACGACGATGACAAGAAGCTTGCCATGGAGCTC; SEQ ID NO: 40), were ligated into the NcoI site (base 88) of HpTK5 in the EcoRV digested Bluescript (Stratagene, La Jolla, CA) vector.

B. In vitro Transcription and Translation

5 Transcription was performed on 2 pmol of linearized HpTK5 or FLAG-HpTK5 containing plasmid at 37°C for 1 h in 50 µl volume containing 10 mM dithiothreitol, 2.5 µg bovine serum albumin, 0.25 mM each dNTP, 0.5 M m7GRNA cap (New England Biolabs, Beverly, MA), 2.5 units RNasin (Promega, Madison, WI), 3 units T3 RNA polymerase (Pharmacia, Piscataway, 10 NJ). 1 µg of DNAase (New England Biolabs, Beverly MA) was added for 15 min at 37°C prior to phenol/chloroform extraction and ethanol precipitation. Translation was performed using the Promega rabbit reticulocyte lysate kit according to the manufacturer's specifications with or without ³⁵S-methionine (350 µCi) labeling. Sample buffer 15 containing SDS and beta-mercaptoethanol (2-ME) was added before boiling and 10% SDS-PAGE.

C. HpTK5 Expression in NIH3T3 Cells

A 4038 bp ClaI - XbaI cDNA fragment containing 32 bp of linker sequence, 37 bp of pBluescript (Stratagene La Jolla, CA) polylinker and 20 the entire 3969 bp HpTK5 cDNA was subcloned into the expression vector pRIS (Genentech, Inc.) under the control of the Rous sarcoma virus LTR promoter. NIH3T3 cells maintained in high glucose Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FCS were co-transfected with pRIS-HpTK5 and pNeo (an SV40 based vector containing the neomycin 25 resistance marker) by the calcium phosphate method as described by Gorman et al., in DNA Prot. Engineer. Tech. 2:3-10 [1990]. Neomycin resistant colonies were selected 48 hours after transfection with Geneticin (Gibco/BRL) at 400 µg/ml. Fourteen days later individual resistant colonies were isolated, expanded and analyzed by flow cytometry for HpTK5 30 expression using rabbit polyclonal antiserum.

D. Immunoprecipitation

Cells (Hep3B, control NIH3T3 or HpTK5 transfected NIH3T3) or in vitro translated protein (HpTK5 or FLAG-HpTK5) were used for immunoprecipitation with either serum (pre-immune or anti-HpTK5-IgG Fc) 35 or monoclonal antibody (FLAG-specific, M2, or isotype control) (IBI,

Rochester, NY). Subconfluent cells were labeled with 200 μ Ci/ml 35 S-methionine for 18 hours and lysed in lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH8.0, 1 mM EDTA, 0.025 Na azide, 1% NP-40, 0.1% SDS, 10% Glycerol, 0.5% Na deoxycholate, 1 mM phenylmethylsulfonyl flouride (PMSF), 10 μ g/ml aprotinin, 10 μ g/ml leupeptin and 50 μ M Na vanadate) for 30 min on ice. The cell lysate was centrifuged (12,000 X g) for 10 min at 4°C. Cell lysate supernatant or in vitro translation mixture was precleared with 0.05 volume of normal rabbit serum and adsorbed with 0.05 volume of Staphylococcus aureus protein-A Sepharose CL4B. After centrifugation, preimmune or immune serum (1:100 dilution), or monoclonal antibody, was added and rocked overnight at 4°C before 100 μ l of protein-A Sepharose CL4B was added and the solution rocked 4°C for additional 2 h. Immunoprecipitates were washed, suspended in SDS/PAGE loading buffer (10% glycerol, 5% 2-ME, 2.3% SDS and 62.5mM Tris-HCl pH 6.8), heated to 95°C for 5 min and analyzed by 7.5% SDS-PAGE.

E. Cell Fractionation

Cell fractionation of Hep3B cells was performed to confirm the membrane localization of HpTK5 predicted by its amino acid sequence. Hep-3B cells (1x10⁷) were labeled with 200 μ Ci/ml 35 S-methionine in alpha MEM medium containing 10% dialyzed FCS overnight. The cells were washed twice with cold PBS, scraped into 1ml of cold buffer (20mM Tris-HCl pH 7.5, 2mM EDTA, 5mM EGTA, 0.25M sucrose, 0.01% leupeptin, 4mM PMSF, 10mM 2-ME) and disrupted by sonication for 40 seconds. Whole homogenates were centrifuged at 12,000 X g for 15 min, the nuclear pellets isolated and the decanted supernatant centrifuged at 140,000 X g for 40 min at 4°C to pellet membranes. The resultant supernatant served as the cytosolic (C) fraction. Nuclear (N) and membrane (M) fractions were washed and dissolved in buffer containing 0.5% NP-40 prior to immunoprecipitation. The C, N or M fractions were immunoprecipitated with an anti-HpTK5 or pre-immune (control) serum, subjected to 12% SDS-PAGE and autoradiographed. HpTK5 segregated predominantly with the membrane fraction, though immunoprecipitated material was evident to a lesser extent in cytosol.

F. Protein Kinase Assay

Immunoprecipitates were washed once with kinase buffer (25mM Hepes pH7.4, 1mM DTT, 10mM MgCl, 10mM MnCl), and resuspended in 40 μ l of kinase

buffer containing either unlabeled ATP or 10 μ Ci of 32 P-ATP (3000Ci/mM). After a 10min incubation at 30°C, the reaction was stopped by adding 40 μ l of 2 X sample buffer and boiling the samples for 3min prior to electrophoresis on 8.0% SDS-PAGE gel. The dried gel was covered with 4
5 sheets of aluminum foil to block 35 S-labelled protein autoradiography and the gel was placed under film for 5 hours to overnight.

G. Western Blotting and Phosphotyrosine Assay

Proteins were electrophoretically transferred to a 0.2 μ m nitrocellulose (Bio-Rad) or a 0.45 μ m polyvinylidene difluoride
10 (Millipore) membrane in a buffer containing 25 mM Tris-HCl (pH 7.5), 192 mM glycine and 20% methanol at 100 mA for 2 h. Filters were washed in TBS (10 mM Tris-HCl pH 8.0, 150 mM NaCl) blocked by incubating in TBST (TBS with 0.05% Tween-20) plus 5% BSA overnight. Filters were washed
15 four times for 5 min each in TBST and incubated for 2 h with 4G10 anti-phosphotyrosine antibody from UBI (1:1000 dilution in TBST). Filters were washed four times for 5 min each in TBST and incubated for 1 h with the alkaline phosphatase labelled anti-mouse secondary antibody (Promega) at a 1:7500 dilution in TBST. After washing four times, the blot was developed for 30-60 min in AP buffer (100mM Tris-HCl, 100 mM NaCl, 5 mM
20 MgCl₂) plus BCIP, NBT substrates.

H. Antibody Induced Phosphorylation Assay

Rabbit antisera to HpTK5-IgG Fc were tested for their ability to induce HpTK5 phosphorylation in HpTK5 transfected NIH3T3 cells. Cells were plated at a density of 5 x 10⁵ cells/well in a 6-well plate and,
25 after 24 hours, were serum starved for 1 hour prior to adding pre-immune or immune serum at a 1:50 dilution for 30 minutes. Cells were then washed in PBS and lysed in either 2X sample buffer or NP-40 lysis buffer as described above. Either crude lysates or immunoprecipitated cell lysates were then separated via 4-12% gradient SDS-PAGE and analyzed by
30 anti-phosphotyrosine immunoblot as described above. HpTK5 expressing cells were exposed to antisera and separated by SDS-PAGE either with or without immunoprecipitation. The electrotransferred gel was immunoblotted with anti-phosphotyrosine antibody. Enhanced tyrosine phosphorylation of HpTK5 was observed following exposure to polyclonal antiserum showing an
35 agonist-like effect of antibody binding. Interaction of HpTK5 with an antibody directed against its ECD induces phosphorylation. This provides

further support that HpTK5 may serve as a receptor for a ligand that triggers kinase activation. Details of the signaling pathway of HpTK5 may be further explored using antisera as a surrogate ligand.

I. Conclusions

5 An HpTK5 ECD-IgG Fc fusion protein was expressed, purified and used to generate rabbit anti-serum which immunoprecipitated a 120kD protein from Hep3B cells. The specificity of the antiserum was confirmed by immunoprecipitation of *in vitro* translated HpTK5 RNA and HpTK5 transfected NIH3T3 cells. To determine the functional capacity of HpTK5, 10 *in vitro* translated HpTK5 was immunoprecipitated, exposed to kinase conditions and immunoblotted using a phosphotyrosine specific monoclonal antibody. The data obtained indicated that HpTK5 is phosphorylated on tyrosine. However, the presence of other bands consistently appearing in the ³²P-labelled immunoprecipitation suggested that HpTK5 protein was 15 only partially purified and therefore, it could not be concluded that HpTK5 was enzymatically active. To overcome this problem, a fusion construct was generated in which an 8 amino acid epitope (FLAG) was added to the N-terminus of HpTK5. The FLAG-HpTK5 fusion was *in vitro* translated and immunoprecipitated with a FLAG-specific monoclonal 20 antibody resulting in a single protein of appropriate size (~120kD). When subjected to kinase conditions in the presence of ³²P-ATP, the HpTK5-FLAG fusion protein was labelled on tyrosine confirming tyrosine autophosphorylation and thereby, the kinase function of HpTK5.

EXAMPLE 5

25 PRODUCTION OF MONOCLONAL ANTIBODIES TO HPTK5

Anti-HpTK5 monoclonal antibodies were produced by hyperimmunizing BALB/c mice intraperitoneally with the HpTK5 extracellular domain (ECD)-human IgG₁ Fc fusion protein (produced using the techniques disclosed above) in RIBI adjuvant (RIBI ImmunoChem Research, Hamilton, MT) and 30 fusing splenocytes with the mouse myeloma cell line X63-Ag8.653 (Kearney et al., J. Immunol. 123:1548-1550 [1979]). The antibodies were purified from ascites fluid using protein A-Sepharose (Repligen Corp., Cambridge, MA) and established affinity chromatography methods (Goding, J.W., J. Immunol. Methods 20:241-253 [1978]).

35 Monoclonal antibodies were screened for their ability to bind the HpTK5 antigen. Starting on day 15 post fusion, culture supernatants were

harvested from the fusion plates and assayed for their ability to specifically "capture" HpTK5-IgG. In this ELISA assay, goat anti-mouse IgG was coated onto 96 well microtiter plates. The culture supernatants (100 μ l) were added to the wells and the mouse IgG present was bound by the goat anti-mouse IgG antibodies. The plates were washed and either HpTK5-IgG or CD4-IgG (100 μ l at 6nM) was added. The "captured" immunoadhesin was detected using a goat anti-hu (Fc specific) horseradish peroxidase conjugate and orthophenylene diamine substrate. Quantitation of substrate catalysis was determined by optical density at 490nm.

Agonist antibodies were then screened for using the techniques disclosed in Example 6 below. Two agonist monoclonal antibodies were identified, one of which has been deposited with the ATCC.

EXAMPLE 6

AGONIST ACTIVITY OF MONOCLONAL ANTIBODIES TO HPTK5

The monoclonal antibodies produced using the techniques disclosed in Example 5 were tested for their ability to induce HpTK5 phosphorylation in HpTK5 transfected NIH3T3 cells. Cells were plated at a density of 5×10^5 cells/well in a 6-well plate and, after 24 hours, were serum starved for 1 hour prior to adding pre-immune serum or anti-HpTK5 monoclonal antibody (undiluted conditioned hybridoma media was used) for 30 minutes. Cells were then washed in PBS and lysed in either 2X sample buffer or NP-40 lysis buffer as described above. Either crude lysates or immunoprecipitated cell lysates were then separated via 4-12% gradient SDS-PAGE and analyzed by anti-phosphotyrosine immunoblot as described above. HpTK5 expressing cells were exposed to the monoclonal antibody and separated by SDS-PAGE either with or without immunoprecipitation. The electrotransferred gel was immunoblotted with anti-phosphotyrosine antibody. Enhanced tyrosine phosphorylation of HpTK5 was observed following exposure to monoclonal antibodies showing an agonist-like effect of antibody binding. Accordingly, interaction of HpTK5 with a monoclonal antibody directed against its ECD is able to induce phosphorylation of the kinase domain thereof.

EXAMPLE 7

PRODUCTION OF POLYCLONAL ANTIBODIES TO SAL-S1

A SAL-S1 extracellular domain (ECD)-human IgG₁ Fc fusion gene was constructed and fusion protein produced as previously described in

Bennett et al., J. Biol. Chem. 266:23060-23067 [1991]. Briefly, PCR primers otk 1.41.1 (SEQ ID NO: 43) and otk 1.41.2 (SEQ ID NO: 44) were employed in the PCR technique using plasmid pRK5.tk1-1.1 (SEQ ID NO: 45) containing SAL-S1 nucleic acid as a template to create a DNA fragment which, when digested with SalI/BstEII, generated an 155bp SalI/BstEII fragment. This 155bp fragment was combined with a 6839bp SalI/HindIII fragment isolated from pRK5.tk1-1.1 and a 719 bp BstEII/HindIII fragment isolated from pBSSK-CH2-CH3 (Bennett et al., *supra*). These fragments were ligated together to create a plasmid pRK5.tk1.ig1.1 (7713bp in size) which, when transfected into 293 cells, was used to produce a SAL-S1 extracellular domain (ECD)-human IgG Fc fusion protein. Fusion protein was prepared and purified as described in Bennett et al., *supra*. Polyclonal antibodies were generated in female New Zealand White rabbits against the fusion protein. Briefly, 12.5µg of fusion protein in 0.625ml PBS was emulsified with 0.625ml Freund's adjuvant (complete adjuvant for the primary injection and incomplete adjuvant for all boosts). The primary injection and all boosts were intramuscular at two sites and subcutaneous at multiple sites. Boosts were carried out at 3 week intervals with bleeds taken 1 and 2 weeks following each boost. SAL-S1 specificity of the immunized rabbit serum was assessed by flow cytometric analysis of 293 (ATCC CRL 1593) and COS7 (ATCC CRL 1651) cells transfected with full length SAL-S1 or vector alone (see below) using a 1:200 dilution of pre-immune serum or anti-SAL-S1-IgG Fc serum. Significant peak shifts were observed in several SAL-S1 expressing clones as compared to either pre-immune serum or vector alone transfectant controls.

EXAMPLE 8

UTILITY AND AGONIST ACTIVITY OF SAL-S1 POLYCLONAL ANTIBODIES

A. Immunoprecipitation

Control 293 and COS7 cells as well as SAL-S1 transfected 293 and COS7 cells were used for immunoprecipitation with either pre-immune serum or anti-SAL-S1-IgG Fc polyclonal antibody. COS7 and 293 cells were transfected using a CaPO₄ procedure as described by Gorman, C. DNA Cloning, Glover D. Ed., IRL Press, Oxford, vol2: 143-190 (1985). For transient expression, 293 cells were transfected as described by Gearing et al. EMBO 8: 3667-3676 (1989). Subconfluent cells were labeled with 200µCi/ml ³⁵S- methionine for 18 hours and lysed in lysis buffer (150 mM

- NaCl, 50mM HEPES, pH 7.5, 1 mM EGTA, 0.025 Na azide, 1% Triton-X 100, 1.5mM MgCl₂, 10% Glycerol, 1 mM phenylmethylsulfonyl flouride [PMSF], 10 µg/ml aprotinin, 10 µg/ml leupeptin and 50 µM Na vanadate) for 10 min on ice. The cell lysate was centrifuged (12,000 X g) for 10 min at 4°C.
- 5 After centrifugation, preimmune or polyclonal antibody was added to the supernatant and rocked for 4 hrs at 4°C before 100 µl of protein-A Sepharose CL4B was added and the solution rocked 4°C for additional 2 h. Immunoprecipitates were washed, suspended in SDS/PAGE loading buffer (10% glycerol, 5% 2-ME, 2.3% SDS and 62.5mM Tris-HCl pH 6.8), heated to 95°C
- 10 for 5 min and analyzed by 7.5% SDS-PAGE.

B. Western Blotting and Phosphotyrosine Assay

- Proteins were electrophoretically transferred to a 0.2 µm nitrocellulose (Bio-Rad) or a 0.45µm polyvinylidene difluoride (Millipore) membrane in a buffer containing 25 mM Tris-HCl (pH 7.5), 192
- 15 mM glycine and 20% methanol at 100 mA for 2 h. Filters were washed in TBS (10 mM Tris-HCl pH 8.0, 150 mM NaCl) blocked by incubating in TBST (TBS with 0.05% Tween-20) plus 5% BSA overnight. Filters were washed four times for 5 min each in TBST and incubated for 2 h with 4G10 anti-phosphotyrosine antibody from UBI (1:1000 dilution in TBST). Filters
- 20 were washed four times for 5 min each in TBST and incubated for 1 h with the alkaline phosphatase labelled anti-mouse secondary antibody (Promega) at a 1:5000 dilution in TBST. After washing four times, the blot was developed for 30-60 min in AP buffer (100mM Tris-HCl, 100 mM NaCl, 5 mM MgCl₂) plus BCIP, NBT substrates.

25 C. Antibody Induced Phosphorylation Assay

- Rabbit antisera to SAL-S1-IgG Fc were tested for their ability to induce SAL-S1 phosphorylation in SAL-S1 transfected 293 cells. Cells were plated at a density of 5 x 10⁵ cells/well in a 6-well plate and, after 24 hours, were serum starved for 12 hours prior to adding pre-immune or
- 30 immune serum at a 1:5 dilution for 30 minutes. Cells were then washed in PBS and lysed in either sample buffer or Triton-X lysis buffer as described above. Either crude lysates or immunoprecipitated cell lysates were then separated via 8% or 4-12% gradient SDS-PAGE and analyzed by anti-phosphotyrosine immunoblot as described above. SAL-S1 expressing
- 35 cells were exposed to antisera and separated by SDS-PAGE either with or without immunoprecipitation. The electrotransferred gel was immunoblotted

with anti-phosphotyrosine antibody. Enhanced tyrosine phosphorylation of SAL-S1 was observed following exposure to polyclonal antiserum showing an agonist-like effect of antibody binding. Interaction of SAL-S1 with an antibody directed against its ECD induces phosphorylation.

5

EXAMPLE 9PRODUCTION OF MONOCLONAL ANTIBODIES TO SAL-S1

Anti-SAL-S1 monoclonal antibodies were produced by hyperimmunizing BALB/c mice in the foot pad with the SAL-S1 extracellular domain-human IgG₁ Fc fusion protein in RIBI adjuvant (RIBI Immunochem Research, Hamilton, MT) and fusing lymphocyte from lymph nodes with the mouse myeloma cell line X63-Ag8U1.

Starting on day 10 post fusion, cultured supernatants were harvest from the fusion plates and assayed for their ability to bind to SAL-S1. In this ELISA assay, SAL-S1 IgG₁ was coated onto 96 microtiter plates. The cultured supernatants (100 μ l) were added to the wells and the mouse antibodies present were bound to Sal-S1 IgG₁. The plates were washed and mouse IgG was detected using a goat anti-mouse IgG (Fc specific with no cross reactivity against human IgG Fc) horseradish peroxidase conjugate and orthophenylene diamine substrate. Quantitation of substrate catalysis was determined by optical density at 490 nm.

Cultured supernatants which were positive from ELISA were then tested for their ability to specifically bind to 293 transfected with SAL-S1 receptor and analyzed by flow cytometry. Agonist antibodies were then screened for using the techniques disclosed in Example 10 below. Six agonist monoclonal antibodies were identified.

EXAMPLE 10AGONIST ACTIVITY OF MONOCLONAL ANTIBODIES TO SAL-S1

The monoclonal antibodies were tested for their ability to induce SAL-S1 phosphorylation in SAL-S1 transfected 293 cells. Cells were harvested from tissue culture dish by assay buffer and washed 2x with the same buffer. 1x10⁵ cells were added to a 96 U-bottom plate which was centrifuged and assay buffer was removed. 150 μ l of cultured supernatants was added to each well followed by incubation at 37°C for 30 minutes, the plate was centrifuged and cultured supernatants were removed. 100 μ l of Fixing solution was added, the cells were fixed for 30 minutes at -20°C, cells were washed with buffer 2x and stained with anti-phosphotyrosine

conjugate with FITC for 60 minutes at 4°C. Cells were analyzed by flow cytometry (FACScan Becton Dickinson, Milpitas, CA). The six anti-SAL-S1 monoclonal antibodies were able to induce SAL-S1 phosphorylation in SAL-S1 transfected 293 cells.

5

Deposit of Materials

The following culture has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Hybridoma</u>	<u>ATCC No.</u>	<u>Deposit Date</u>
Anti-HpTK5	HB 11,583	March 15, 1994

10 This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture for 30 years from the date of deposit. The organism will be made available by ATCC under
15 the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures
20 availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if the
25 culture on deposit should die or be lost or destroyed when cultivated under suitable conditions, it will be promptly replaced on notification with a viable specimen of the same culture. Availability of the deposited strain is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of
30 any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the culture deposited, since the deposited embodiment is intended as a single illustration of one
35 aspect of the invention and any culture that are functionally equivalent

are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as
5 limiting the scope of the claims to the specific illustration that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

10

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Genentech, Inc.
Bennett, Brian D.
5 Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
- 10 (ii) TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
- (iii) NUMBER OF SEQUENCES: 45
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 460 Point San Bruno Blvd
15 (C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080
- 20 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: patin (Genentech)
- 25 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- 30 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/222616
(B) FILING DATE: 04-APR-1994
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Wendy M. Lee
(B) REGISTRATION NUMBER: 00,000
(C) REFERENCE/DOCKET NUMBER: 821P3PCT
- 35 (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415/225-1994
(B) TELEFAX: 415/952-9881
(C) TELEX: 910/371-7168
- (2) INFORMATION FOR SEQ ID NO:1:
- 40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGGATCCACA GNGACCT 17

(2) INFORMATION FOR SEQ ID NO:2:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10 GGAATTCCAA AGGACCAGAC GTC 23

(2) INFORMATION FOR SEQ ID NO:3:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGGATCCATC CACAGAGATG T 21

(2) INFORMATION FOR SEQ ID NO:4:

- 20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGAATTCCTT CAGGAGCCAT CCACTT 26

(2) INFORMATION FOR SEQ ID NO:5:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 160 bases
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGATCCTGTG CATCAGTGAC TTAGGGCTAG GAACATTCTG CTGTCGGAAA 50
5 GCGACGTGGT GAAGATCTGT GACTTTGGCC TTGCCCGGGA CATCTACAAA 100
GACCCCAGCT ACGTCCGCAA GCATGCCCGG CTGCCCCTGA AGTGGATGGC 150
GCCAGAATTC 160

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
10 (A) LENGTH: 53 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Pro Val His Gln Xaa Leu Arg Ala Arg Asn Ile Leu Leu Ser
15 1 5 10 15
Glu Ser Asp Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp
20 25 30
Ile Tyr Lys Asp Pro Ser Tyr Val Arg Lys His Ala Arg Leu Pro
35 40 45
20 Leu Lys Trp Met Ala Pro Glu Phe
50 53

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 147 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGATCCATTC ACAGAGACCT AGCAGCACGC AACATCCTGG TCTCAGAGGA 50
30 CCTGGTAACC AAGGTCAGCG ACTTTGGCCT GGCCAAAGCC GAGCGGAAGG 100

GGCTAGACTC AAGCCGGCTG CCCGTCAAAT GGATGGCTCC CGAATTC 147

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Ser Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ser
1 5 10 15
10 Glu Asp Leu Val Thr Lys Val Ser Asp Phe Gly Leu Ala Lys Ala
20 25 30
Glu Arg Lys Gly Leu Asp Ser Ser Arg Leu Pro Val Lys Trp Met
35 40 45
Ala Pro Glu Phe
15 49

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 149 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTTGGAATTC CTTCCGGCGC CATCCATTTT ACCGGCAGCT TTATTTTCGTG 50
TCTAGATTCA TAGATGTCTT CATTATCTAC CTAAAAACT CTGGCAAGTC 100
25 CAAAATCTGC TACTTTGTAG ATATTATGTT CACCAACGAG GACATTCCT 149

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Gly Ile Pro Ser Gly Ala Ile His Phe Thr Gly Ser Phe Ile
1 5 10 15

Ser Cys Leu Asp Ser Met Ser Ser Leu Ser Thr Leu Lys Thr Leu
 20 25 30

Ala Ser Pro Lys Ser Ala Thr Leu Ile Leu Cys Ser Pro Thr Arg
 35 40 45

5 Thr Phe
 47

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 151 bases
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GTGCACAGGG ATCTCGCGGC TCGGAACATC CTCGTCGGGG AAAACACCCT 50

15 CTCGAAAGTT GGGGACTTCG GGTTAGCCAG GCTTATCAAG GAGGACGTCT 100

ACCTCTCCCA TGACCACAAT ATCCCCTACA AATGGATGGC CCCTGAGGGA 150

A 151

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 50 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

25 Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Gly Glu Asn
 1 5 10 15

Thr Leu Ser Lys Val Gly Asp Phe Gly Leu Ala Arg Leu Ile Lys
 20 25 30

Glu Asp Val Tyr Leu Ser His Asp His Asn Ile Pro Tyr Lys Trp
 35 40 45

30 Met Ala Pro Glu Gly
 50

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 137 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

G TTCACCGAG ATCTCAAGTC CAACAACATT TTGCTGCTGC AGCCCATTGA 50

GAGTGACGAC ATGGAGCACA AGACCCTGAA GATCACCGAC TTTGGCCTGG 100

CCCGAGAGTG GCACAAAACC ACACAAATGA GTGCCGC 137

(2) INFORMATION FOR SEQ ID NO:14:

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

15	Val	His	Arg	Asp	Leu	Lys	Ser	Asn	Asn	Ile	Leu	Leu	Leu	Gln	Pro
	1				5					10				15	
	Ile	Glu	Ser	Asp	Asp	Met	Glu	His	Lys	Thr	Leu	Lys	Ile	Thr	Asp
					20					25				30	
20	Phe	Gly	Leu	Ala	Arg	Glu	Trp	His	Lys	Thr	Thr	Gln	Met	Ser	Ala
					35					40				45	

(2) INFORMATION FOR SEQ ID NO:15:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 211 bases
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCAATCGTG ACCTCGCCGC CCGAAATGTG TTGCTAGTTA CCCAACATTA 50

CGCCAAGATC AGTGATTTTCG GACTTTCCAA AGCACTGCGT GCTGATGAAA 100

30 ACTACTACAA GGCCCAGACC CATGGAAAGT GGCCTGTCAA GTGGTACGCT 150

CCGGAATGCA TCAACTACTA CAAGTTCTCC AGCAAAAGCG ATGTCTGGTC 200

CTTTGGAATT C 211

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 70 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

10	Val	Asn	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu	Leu	Val	Thr	Gln
	1				5					10					15
	His	Tyr	Ala	Lys	Ile	Ser	Asp	Phe	Gly	Leu	Ser	Lys	Ala	Leu	Arg
					20					25					30
	Ala	Asp	Glu	Asn	Tyr	Tyr	Lys	Ala	Gln	Thr	His	Gly	Lys	Trp	Pro
					35					40					45
15	Val	Lys	Trp	Tyr	Ala	Pro	Glu	Cys	Ile	Asn	Tyr	Tyr	Lys	Phe	Ser
					50					55					60
	Ser	Lys	Ser	Asp	Val	Trp	Ser	Phe	Gly	Ile					
					65					70					

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 6827 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 50

TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC 100

TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG 150

ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA 200

30 TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCAC TTGGCAGTAC 250

ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300

AAATGGCCCG CCTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC 350

TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC 400

GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 450

5 TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA 500

AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC 550

AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600

TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT 650

CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA 700

10 TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA 750

GTCTATAGGC CCACTTGGCT TCGTTAGAAC GCGGCTACAA TTAATACATA 800

ACCTTATGTA TCATACACAT ACGATTTAGG TGACACTATA GAATAACATC 850

CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900

CGGTTCTATC GATTGAATTC CCCGGGGATC CTCTAGAGAT CCCTCGACCT 950

15 CGAGATCCAT TGTGCTGGCG CGGATTCTTT ATCACTGATA AGTTGGTGGA 1000

CATATTATGT TTATCAGTGA TAAAGTGTCA AGCATGACAA AGTTGCAGCC 1050

GAATACAGTG ATCCGTGCCG CCCTAGACCT GTTGAACGAG GTCGGCGTAG 1100

ACGGTCTGAC GACACGCAAA CTGGCGGAAC GGTGGGGGGT TCAGCAGCCG 1150

GCGCTTTACT GGCACCTCAG GAACAAGCGG GCGCTGCTCG ACGCACTGGC 1200

CGAAGCCATG CTGGCGGAGA ATCATAGCAC TTCGGTGCCG AGAGCCGACG 1250

ACGACTGGCG CTCATTTCTG ACTGGGAATG CCCGCAGCTT CAGGCAGGCG 1300

CTGCTCGCCT ACCGCCAGCA CAATGGATCT CGAGGGATCT TCCATACCTA 1350

CCAGTTCTGC GCCTGCAGGT CGCGGCCGCA CTACTCTTTG ATGTATTACT 1400

5 CATATTACCA AGGAATAACT GGCGGGCACA GGGTCAGGTG CTGAAGGGAC 1450

ATTGTGAGAA GTGACCTAGA AGGCAAGAGG TGAGCCCTCT GTCACGCTGG 1500

CATAAGGGCC GCTTGAGGGC TCTTTGGTCA AGCAGTAACG CCAGTGTCTG 1550

GGAAGGCACC TGTTACTCAG CAGACCATGA AAGGGCGTCT CCCTTTCCTT 1600

GGAGCAGTCA GGAACACTC TGCTCCACCA GCTTCTTGTG GGAGCCTGGA 1650

10 TATTATCCAG GCCTGCCCGC AGTCATCCGG AGGCCTAACC CCTCCCTGTG 1700

GTGCTTCAGT GGTCACTC CTTGTCCACT TTCATGCTCC TCTTGGCCTC 1750

CTGGTTCCTC TTGGAAGTTT GTAGTAGATA GCAGAAGAAA TAGCGAAAGT 1800

CTTAAAGTCT TTGATCTTTC TTATAAGTGC AGAGAAGAAA TGCTGACGTA 1850

TGCTGCCTTC TCTCTCTCTG CTTCAGCTAC CTGAAGCCGC TTTCTTGTCT 1900

15 ATACCTGCTC TCTATCTGCT CACACTCCTC CGAGGCCAGC ACCATCCCAC 1950

TGTCTGTCTG GTTGTCCACA GAGCCTTTGT AGGTCGTTGG GGTCATGGGG 2000

AATTCCTCAA ATGTCTTCAT CCTGGAGGAA CCACGGGTCT CAGCCCCTCT 2050

GGCCAGGCAC CCGGGAAAGG ACACCCAGTT GTAATACCTG GCGGCCAGGC 2100

TGTGGCGCTG CAGGCTTGGC GGGCTGTCCT CAGCGTCAGC CTGGGCGATG 2150

TGTAGGGCCA TGGTGGACAC CTGCGAGAAG CTGCCCTCTT CTGAGCTCTG 2200

AGAGCTGCGC GGGGCCATGC AGACCTCCTC TTCCTCTTGC AGGCCCCTGC 2250

CCTGGAGCAG GTCCCCCAGG ATCTCCACCA GCTCCGAGAA TGCAGGTCTC 2300

GCCTTGGGGT CTCCGGACCA GCAGTTCAGC ATGATGCGGC GTATGGCGGG 2350

5 AGTGGCCAGC TCCGGGGCCC TCATCCTTGT GCCGTCTCTC AGCCGCTGGC 2400

AGAACTCCTC ATTGATCTGC ACCCCAGGGT ACGGGGAGGC CCCCAGAGAG 2450

AAGATCTCCC AGAGAAGCAC CCCAAAGGAC CACACGTCAC TCTGCGTGGT 2500

GTACACCTTG TCGAAGATGC TTTCAGGGGC CATCCACTTC AGGGGCAGCC 2550

GGGCACTGCC CTTGCGGACG TAGTCGGGGT CTTTGTAGAT GTCCCCGGGCA 2600

10 AGGCCAAAGT CACAGATCTT CACCACGTCG CTTTCCGACA GCAGAATGTT 2650

CCGAGCAGCC AGGTCTCTGT GGATGCACTT TCGGGAAGCC AGGAACTCCA 2700

TCCCTCTGGC CACCTGGAAG CTGTAGCAGA CAAGATCTTC CATGGTCAGC 2750

GGGCTCAGCC ACAGGTCCTC AGCTTCTTGG TCTGGAGAAG CCCGCCTCGC 2800

TCCGCCCTCG GTCTTCGAGA ACCGCGCGAA GAGGACCCTG TCGCTGCTCC 2850

15 CCGGCCGCTT CCGATCCAGC CTGGCGAGCT CCACCATGGC GCGGAAGCGT 2900

CCGCGCTGCT CGGGAGACTT CTCCTGCGGA TGCACGAAGC TGGCTCGAGG 2950

GCGCCCAGTC GTCCGCCGCA GAGGCGCCTC CATTCCCCCG CCGCCGCGG 3000

CGCCCCGAG GCCGCCGCT CACCGNGCAG GGGCTGCGGC CGCGACTCTA 3050

GAGTCGACCT GCAGAAGCTT GGCCGCCATG GCCCAACTTG TTTATTGCAG 3100

CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA 3150

GCATTTTTTTT CACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT 3200

ATCTTATCAT GTCTGGATCG ATCGGGAATT AATTCGGCGC AGCACCATGG 3250

CCTGAAATAA CCTCTGAAAG AGGAACTTGG TTAGGTACCT TCTGAGGCGG 3300

5 AAAGAACCAG CTGTGGAATG TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG 3350

GCTCCCCAGC AGGCAGAAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA 3400

ACCAGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA GTATGCAAAG 3450

CATGCATCTC AATTAGTCAG CAACCATAGT CCGCCCCCTA ACTCCGCCCA 3500

TCCCGCCCCCT AACTCCGCCC AGTTCCGCCC ATTCTCCGCC CCATGGCTGA 3550

10 CTAATTTTTT TTATTTATGC AGAGGCCGAG GCCGCCTCGG CCTCTGAGCT 3600

ATTCCAGAAG TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC TTTTGCAAAA 3650

AGCTGTTAAC AGCTTGGCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG 3700

AAAACCCTGG CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCCTTC 3750

GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC CTTCCCAACA 3800

15 GTTGCGTAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT TTTCTCCTTA 3850

CGCATCTGTG CGGTATTTCA CACCGCATAC GTCAAAGCAA CCATAGTACG 3900

CGCCCTGTAG CGGCGCATT AAGCGCGCGG GTGTGGTGGT TACGCGCAGC 3950

GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCGCTCCTT TCGCTTTCTT 4000

CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 4050

GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC 4100

AAAAAATTG ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA 4150

GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC 4200

TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGG CTATTCTTTT 4250

5 GATTTATAAG GGATTTTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT 4300

GATTTAACA AAATTTAACG CGAATTTTAA CAAAATATTA ACGTTTACAA 4350

TTTTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC 4400

CAACTCCGCT ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG 4450

CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC 4500

10 TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT 4550

CACCGTCATC ACCGAAACGC GCGAGGCAGT ATTCTTGAAG ACGAAAGGGC 4600

CTCGTGATAC GCCTATTTTT ATAGGTTAAT GTCATGATAA TAATGGTTTC 4650

TTAGACGTCA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT 4700

GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA 4750

15 CCCTGATAAA TCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA 4800

ACATTTCCGT GTCGCCCTTA TTCCCTTTTT GCGGCATT TGCCTTCCTG 4850

TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG 4900

TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT 4950

CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA 5000

AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTGATGACGC CGGGCAAGAG 5050

CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC 5100

ACCAATCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT 5150

GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG 5200

5 ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG 5250

GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA 5300

TACCAAACGA CGAGCGTGAC ACCACGATGC CAGCAGCAAT GGCAACAACG 5350

TTGCGCAAAC TATTAAGTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA 5400

ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT 5450

10 CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG 5500

CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC 5550

CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC 5600

GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA 5650

CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA 5700

15 TTTTAAATTT AAAAGGATCT AGGTGAAGAT CCTTTTGTAT AATCTCATGA 5750

CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA 5800

GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG 5850

CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG 5900

ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG 5950

CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT 6000

CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC 6050

CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA 6100

AGACGATAGT TACCGGATAA GGC GCAGCGG TCGGGCTGAA CCGGGGGTTC 6150

5 GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC 6200

TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG 6250

GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA 6300

GCTTCCAGGG GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC 6350

ACCTCTGACT TGAGCGTCGA TTTTGTGAT GCTCGTCAGG GGGGCGGAGC 6400

10 CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG 6450

CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG 6500

ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA 6550

ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT 6600

ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TCCAGCTGGC 6650

15 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT 6700

GTGAGTTACC TCACTCATTG GGCACCCAG GCTTTACACT TTATGCTTCC 6750

GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA 6800

ACAGCTATGA CCATGATTAC GAATTAA 6827

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 348 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

	Glu	Lys	Ser	Pro	Glu	Gln	Arg	Gly	Arg	Phe	Arg	Ala	Met	Val	Glu	
	1				5					10					15	
	Leu	Ala	Arg	Leu	Asp	Arg	Arg	Arg	Pro	Gly	Ser	Ser	Asp	Arg	Val	
					20					25					30	
10	Leu	Phe	Ala	Arg	Phe	Ser	Lys	Thr	Glu	Gly	Gly	Ala	Arg	Arg	Ala	
					35					40					45	
	Ser	Pro	Asp	Gln	Glu	Ala	Glu	Asp	Leu	Trp	Leu	Ser	Pro	Leu	Thr	
					50					55					60	
	Met	Glu	Asp	Leu	Val	Cys	Tyr	Ser	Phe	Gln	Val	Ala	Arg	Gly	Met	
15					65					70					75	
	Glu	Phe	Leu	Ala	Ser	Arg	Lys	Cys	Ile	His	Arg	Asp	Leu	Ala	Ala	
					80					85					90	
	Arg	Asn	Ile	Leu	Leu	Ser	Glu	Ser	Asp	Val	Val	Lys	Ile	Cys	Asp	
					95					100					105	
20	Phe	Gly	Leu	Ala	Arg	Asp	Ile	Tyr	Lys	Asp	Pro	Asp	Tyr	Val	Arg	
					110					115					120	
	Lys	Gly	Ser	Ala	Arg	Leu	Pro	Leu	Lys	Trp	Met	Ala	Pro	Glu	Ser	
					125					130					135	
	Ile	Phe	Asp	Lys	Val	Tyr	Thr	Thr	Gln	Ser	Asp	Val	Trp	Ser	Phe	
25					140					145					150	
	Gly	Val	Leu	Leu	Trp	Glu	Ile	Phe	Ser	Leu	Gly	Ala	Ser	Pro	Tyr	
					155					160					165	
	Pro	Gly	Val	Gln	Ile	Asn	Glu	Glu	Phe	Cys	Gln	Arg	Leu	Arg	Asp	
					170					175					180	
30	Gly	Thr	Arg	Met	Arg	Ala	Pro	Glu	Leu	Ala	Thr	Pro	Ala	Ile	Arg	
					185					190					195	
	Arg	Ile	Met	Leu	Asn	Cys	Trp	Ser	Gly	Asp	Pro	Lys	Ala	Arg	Pro	
					200					205					210	
	Ala	Phe	Ser	Glu	Leu	Val	Glu	Ile	Leu	Gly	Asp	Leu	Leu	Gln	Gly	
35					215					220					225	
	Arg	Gly	Leu	Gln	Glu	Glu	Glu	Glu	Val	Cys	Met	Ala	Pro	Arg	Ser	
					230					235					240	
	Ser	Gln	Ser	Ser	Glu	Glu	Gly	Ser	Phe	Ser	Gln	Val	Ser	Thr	Met	
					245					250					255	

Ala Leu His Ile Ala Gln Ala Asp Ala Glu Asp Ser Pro Pro Ser
 260 265 270

Leu Gln Arg His Ser Leu Ala Ala Arg Tyr Tyr Asn Trp Val Ser
 275 280 285

5 Phe Pro Gly Cys Leu Ala Arg Gly Ala Glu Thr Arg Gly Ser Ser
 290 295 300

Arg Met Lys Thr Phe Glu Glu Phe Pro Met Thr Pro Thr Thr Tyr
 305 310 315

10 Lys Gly Ser Val Asp Asn Gln Thr Asp Ser Gly Met Val Leu Ala
 320 325 330

Ser Glu Glu Cys Glu Gln Ile Glu Ser Arg Tyr Arg Gln Glu Ser
 335 340 345

Gly Phe Arg
 348

15 (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7607 bases
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 50

TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCGCG GTTACATAAC 100

TTACGGTAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG 150

25 ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA 200

TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC TTGGCAGTAC 250

ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300

AAATGGCCCG CCTGGCATTG TGCCCACTAC ATGACCTTAT GGGACTTTCC 350

TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC 400

30 GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 450

TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA 500

AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC 550

AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600

TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT 650

5 CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA 700

TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA 750

GTCTATAGGC CCACTTGGCT TCGTTAGAAC GCGGCTACAA TTAATACATA 800

ACCTTATGTA TCATACACAT ACGATTTAGG TGACACTATA GAATAACATC 850

CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900

10 CGGTTCTATC GATTGAATTC CCCGGGGATC CTCTAGAGAT CCCTCGACCT 950

CGAGTCGACT TTTTTTTTTT TTTTGTAGG CCAAAGGGTA CTTCTTTTTC 1000

TTTATTAATT ACTCAGAAGT CTAGGCCACA GCAATCTACT GTTCTCCTCT 1050

CATTTTCCTA AACTATTTTG ATACCTATTT CTCAGACTTT ATGGGCTATT 1100

AGACATTTCT CACATTTCCA TAGATAATAA CTCATCCGTT TTGCAACCTG 1150

15 ATTCTCAATA TTAAGAGATT AAACTAATG TATATGACTC TCAGTTGACA 1200

CATACTGAAG TACAGAAAAA TTCCATCATT TCCTTCTGCA AAATGAAAAA 1250

GACTTCGTTT TCTCAACAGC TGCATCATTT TTTTATGCAT AGAAAAAAT 1300

GTGCAATTAC TCCAAGTACA ATCAAGTCAT TTAACATGGC TTTACCATCA 1350

TTGTAGTTAC AGGATATTTT AAAAGAGAAA AAAAAATCTC AAAGCACAGG 1400

TCCTGCTGTG CAGCAAAGCA ATCAAATTCC TTCATAATAA CAGCCTGATG 1450

GGATTCAGCA ATCTGAGGAA TAATGAATAA CCACTCTAAT CAGTAAACAG 1500

GAAAATGCTA CAACAGTCAC TGAGTAAAAA TTGGACTATC ATCTGTTGAT 1550

TCTCTTGATC GACATTTCAA ACAATAAATG GAAATGTAAG TATCTCTTAA 1600

5 AAAGAAAAAT AACTTGTTTT AGTGTGCTTA ATTTTACCAG GCAGTGAGGA 1650

AATTATATAT CACCTTGACT GTCCTGCAGT GTTGCCCAGT CAATAAAATG 1700

CACAAATAAT CTTTTTCATA ATACATGGCC AACTTTATCC TATCACTTGA 1750

ATATGTCAGG ATAAACTGAT TGTGCAGTTG GTTGATAACA TTGTATTTTG 1800

GAATGGATTA TTTGAATTTG TTTTGCTACT TTATTATTTG ATATTCTTCT 1850

10 CCAGTGTTCA TCTTATGAAG TTATTTGCAT CTGAATATGA AGAGTCTGTT 1900

TCAAAATAGT CTTCAAGTTT CCAACGCAGT GTCTCAAATG TAGGTCGTTT 1950

CTTAGGCTCT GCATTCCAGC ACTCCAACAT GATGTTGTAA AATTGCTGTG 2000

GACAGTTGGA TGGTTGCGGA AGTCTATAGT TTTGAGCCAA CATCTGGATT 2050

ACCTGGGCAC CTGTCATACC ACTGTAAGGC ATTTTGCCAT AAGTAATGAT 2100

15 TTCATAAAGA AGGATTCCAA ATGACCATAC ATCGGACTTA ATGCTGAATT 2150

TATTACTACG AATGGCTTCG GGCAGTCC ACTTCACCGG CAGCTTTATT 2200

TCGTGTCTAG ATTCATAGAT GTCTTCATTA TCTACCTTAA AACTCTGGC 2250

AAGTCCAAAA TCTGCTACTT TGTAGATATT ATGTTACCA ACGAGGACAT 2300

TTCTGGCAGC CAGATCTCTG TGAATGTAGT TCCGAGACTC CAGATAGGCC 2350

ATTCCAGAGG CAACCTGTGC CGCCATGTCT ACCTGTTGAG TCAGATGGAT 2400

TTTGTATCCA GTGTCATTTT GGAGATATTC TTGCAGACTT CCATGTCTCA 2450

TCAACTCTGT AATAATATAA ATTGGATCTT CTAAAGTGCA AACAGCATAA 2500

AGCTGGATAA GCTTTGGATG TCTTAGGTTC TTCATTATCT GTGCCTCCCT 2550

5 CAGGAAGTCA TTTGGATCCA TTGAACCTGG TTTTAATGTT TTTACTGCTA 2600

CTGGAGTGGT ATTGTTCCAC AGACCTTCCC ATACTTCGCC AACTGACCA 2650

GATCCCAATC GCTTCAGAAG CTGTATGGAG TTGCGGTCTA TCTCCCATTG 2700

GTCCACGGTT TTATACGACA AATCAAATGG AGCTGGGACC TGGATCTTTA 2750

AGCATGGTTT CCCAGCTTG ACACACAGGC CGTCACTTGT CTTGGTGTAG 2800

10 TGGCTCACAA ATTCGTTTCTG TGTGAAAAG ATTCTTCTTC GCGTGAGAAA 2850

AAATCCCCCT TCATCCAGTC TTTTAATTCT GTAGTGTCTT ACAACTGCTC 2900

CATCTAAAC TGAAAGAGAG AATTCTCCTT TTTGGCTTTC ACTTTCTCTG 2950

ATTAGAAAGG AACCGGTCTT GTTTTCTGAA TATAATAGTT GTTTCTCTGC 3000

ATCTGATCTT CCGATTGCTC CAAAGAACCA CGGCTCTGCC TGTAGGCTTC 3050

15 TGTCTCAGC CACGTAGTTA GAAGGAATAT AGCCTTGTAG TTGCTGACTG 3100

GAGCCATCTC GTCTTTTCTC CAAGTGTCTG GCAAACCACC AGCCCTCATG 3150

CAAAGTGTCC AGAACTTGAA GTTTGTCAAC TGCTCGGAAG CTCAAGTCCT 3200

CAGCAGTCCG AGCCTGGTAA TCAAACAAAG CCACAAAGTA GTGGCCATGC 3250

CTCTGTGACT GGGGAGAGCA AAGGGCCCCCT GGATTTTCAA TCACGGTTGA 3300

CTTGTCTGCC TCCGTGGACA AACAGGGGAG ATAGGGTTCT AGGTACTCCC 3350

AGAGCCTCTG ACAGATGTTG CTCATTGTGC CTTGGTGGGG AGAAGAGGAG 3400

CAGGGCTTCT CCCTCTCCCC TTAGTCTCTG CGATCCACCT TATCTTCCTT 3450

CACCAGGCAA CTTTGAAGTC AGCACCAACT CACCATACTT CGGAGAGTAT 3500

5 GCAAAGTCCC GTTTCAGATC AGTCCAGCAG CTGGGTTGCA GCAAGTCCTA 3550

CCTGGAGAGA CTTACCGGCT TGCTTTCTGT GGCTGGAGGT GCTACCCCGA 3600

GGCAAACTG AGCAGGAGCT GGGCAGCTGC TCACTAGGAA GGTGTCTTTT 3650

CTTCTTATCT GCTTAAGAAT CCCACAACAA AAATAAAATA AAATTAAAAG 3700

GGCTTTATTT AGACAAATAT CTGAGAACAG AATGGTGCCA TCTTGCCTTT 3750

10 TGTCCCAATA AAAAGTTAGC AAGAGGAAGC TACTAACCCC TGGTAAAACC 3800

TCCACGTCTT GCTTTCGCCA GGGTCGACTC GAGGGATCTT CCATACCTAC 3850

CAGTTCTGCG CCTGCAGGTC GCGGCCGCGA CTCTAGAGTC GACCTGCAGA 3900

AGCTTGGCCG CCATGGCCCA ACTTGTTTAT TGCAGCTTAT AATGGTTACA 3950

AATAAAGCAA TAGCATCACA AATTCACAA ATAAAGCATT TTTTTCCTG 4000

15 CATTCTAGTT GTGGTTTGTC CAACTCATC AATGTATCTT ATCATGTCTG 4050

GATCGGGAAT TAATTCGGCG CAGCACCATG GCCTGAAATA ACCTCTGAAA 4100

GAGGAACTTG GTTAGGTACC TTCTGAGGCG GAAAGAACCA GCTGTGGAAT 4150

GTGTGTCAGT TAGGGTGTGG AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG 4200

TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAAGTCCC 4250

CAGGCTCCCC AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA 4300

GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC 4350

CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG 4400

CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA 4450

5 GGCTTTTTTG GAGGCCTAGG CTTTGTGCAAA AAGCTGTTAA CAGCTTGGCA 4500

CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA 4550

ACTTAATCGC CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG 4600

AAGAGGCCCC CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC 4650

GAATGGCGCC TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTT 4700

10 ACACCGCATA CGTCAAAGCA ACCATAGTAC GCGCCCTGTA GCGGCGCATT 4750

AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT ACACTTGCCA 4800

GCGCCCTAGC GCCCGCTCCT TTCGCTTTCT TCCCTTCCTT TCTCGCCACG 4850

TTCGCCGGCT TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC CTTTAGGGTT 4900

CCGATTTAGT GCTTTACGGC ACCTCGACCC CAAAAAATT GATTTGGGTG 4950

15 ATGGTTCACG TAGTGGGCCA TCGCCCTGAT AGACGGTTTT TCGCCCTTTG 5000

ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC AACTGGAAC 5050

AACACTCAAC CCTATCTCGG GCTATTCTTT TGATTATAA GGGATTTTGC 5100

CGATTTCCGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC 5150

GCGAATTTTA ACAAATATT AACGTTTACA ATTTTATGGT GCACTCTCAG 5200

TACAACTGTC TCTGATGCCG CATAGTTAAG CCAGCCCCGA CACCCGCCAA 5250

CACCCGCTGA CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC 5300

AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA GGTTTTCCACC 5350

GTCATCACCG AAACGCGCGA GACGAAAGGG CCTCGTGATA CGCCTATTTT 5400

5 TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT 5450

TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA 5500

TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT 5550

AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT 5600

ATTCCCTTTT TTGCGGCAAT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC 5650

10 GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT 5700

ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC 5750

GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC 5800

GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC 5850

ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT 5900

15 CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT 5950

GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA 6000

AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT 6050

GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA 6100

CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGTG 6150

CGGAAGTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG 6200

GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG 6250

GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA 6300

TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC 6350

5 ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA 6400

GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT 6450

CATATATACT TTAGATTGAT TTAAAACTTC ATTTTAAATT TAAAAGGATC 6500

TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA 6550

GTTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT 6600

10 CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAA 6650

CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT 6700

TTTTCCGAAG GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATACTGTTC 6750

TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG 6800

CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG 6850

15 CGATAAGTCG TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA 6900

AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACA GCCCAGCTTG 6950

GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA 7000

AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG 7050

GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC 7100

TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG 7150
 ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA 7200
 ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG 7250
 TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT 7300
 5 TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT 7350
 CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC 7400
 GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG 7450
 GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT 7500
 AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA 7550
 10 ATGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACATGATTAC 7600
 GAATTAA 7607

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 505 amino acids
 15 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	Met	Ser	Asn	Ile	Cys	Gln	Arg	Leu	Trp	Glu	Tyr	Leu	Glu	Pro	Tyr
	1				5					10					15
20	Leu	Pro	Cys	Leu	Ser	Thr	Glu	Ala	Asp	Lys	Ser	Thr	Val	Ile	Glu
					20					25					30
	Asn	Pro	Gly	Ala	Leu	Cys	Ser	Pro	Gln	Ser	Gln	Arg	His	Gly	His
					35					40					45
25	Tyr	Phe	Val	Ala	Leu	Phe	Asp	Tyr	Gln	Ala	Arg	Thr	Ala	Glu	Asp
					50					55					60
	Leu	Ser	Phe	Arg	Ala	Gly	Asp	Lys	Leu	Gln	Val	Leu	Asp	Thr	Leu
					65					70					75

	His	Glu	Gly	Trp	Trp	Phe	Ala	Arg	His	Leu	Glu	Lys	Arg	Arg	Asp	80	85	90
	Gly	Ser	Ser	Gln	Gln	Leu	Gln	Gly	Tyr	Ile	Pro	Ser	Asn	Tyr	Val	95	100	105
5	Ala	Glu	Asp	Arg	Ser	Leu	Gln	Ala	Glu	Pro	Trp	Phe	Phe	Gly	Ala	110	115	120
	Ile	Gly	Arg	Ser	Asp	Ala	Glu	Lys	Gln	Leu	Leu	Tyr	Ser	Glu	Asn	125	130	135
10	Lys	Thr	Gly	Ser	Phe	Leu	Ile	Arg	Glu	Ser	Glu	Ser	Gln	Lys	Gly	140	145	150
	Glu	Phe	Ser	Leu	Ser	Val	Leu	Asp	Gly	Ala	Val	Val	Lys	His	Tyr	155	160	165
	Arg	Ile	Lys	Arg	Leu	Asp	Glu	Gly	Gly	Phe	Phe	Leu	Thr	Arg	Arg	170	175	180
15	Arg	Ile	Phe	Ser	Thr	Leu	Asn	Glu	Phe	Val	Ser	His	Tyr	Thr	Lys	185	190	195
	Thr	Ser	Asp	Gly	Leu	Cys	Val	Lys	Leu	Gly	Lys	Pro	Cys	Leu	Lys	200	205	210
20	Ile	Gln	Val	Pro	Ala	Pro	Phe	Asp	Leu	Ser	Tyr	Lys	Thr	Val	Asp	215	220	225
	Gln	Trp	Glu	Ile	Asp	Arg	Asn	Ser	Ile	Gln	Leu	Leu	Lys	Arg	Leu	230	235	240
	Gly	Ser	Gly	Gln	Phe	Gly	Glu	Val	Trp	Glu	Gly	Leu	Trp	Asn	Asn	245	250	255
25	Thr	Thr	Pro	Val	Ala	Val	Lys	Thr	Leu	Lys	Pro	Gly	Ser	Met	Asp	260	265	270
	Pro	Asn	Asp	Phe	Leu	Arg	Glu	Ala	Gln	Ile	Met	Lys	Asn	Leu	Arg	275	280	285
30	His	Pro	Lys	Leu	Ile	Gln	Leu	Tyr	Ala	Val	Cys	Thr	Leu	Glu	Asp	290	295	300
	Pro	Ile	Tyr	Ile	Ile	Thr	Glu	Leu	Met	Arg	His	Gly	Ser	Leu	Gln	305	310	315
	Glu	Tyr	Leu	Gln	Asn	Asp	Thr	Gly	Ser	Lys	Ile	His	Leu	Thr	Gln	320	325	330
35	Gln	Val	Asp	Met	Ala	Ala	Gln	Val	Ala	Ser	Gly	Met	Ala	Tyr	Leu	335	340	345
	Glu	Ser	Arg	Asn	Tyr	Ile	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	350	355	360

Leu Val Gly Glu His Asn Ile Tyr Lys Val Ala Asp Phe Gly Leu
 365 370 375
 Ala Arg Val Phe Lys Val Asp Asn Glu Asp Ile Tyr Glu Ser Arg
 380 385 390
 5 His Glu Ile Lys Leu Pro Val Lys Trp Thr Ala Pro Glu Ala Ile
 395 400 405
 Arg Ser Asn Lys Phe Ser Ile Lys Ser Asp Val Trp Ser Phe Gly
 410 415 420
 10 Ile Leu Leu Tyr Glu Ile Ile Thr Tyr Gly Lys Met Pro Tyr Ser
 425 430 435
 Gly Met Thr Gly Ala Gln Val Ile Gln Met Leu Ala Gln Asn Tyr
 440 445 450
 Arg Leu Pro Gln Pro Ser Asn Cys Pro Gln Gln Phe Tyr Asn Ile
 455 460 465
 15 Met Leu Glu Cys Trp Asn Ala Glu Pro Lys Glu Arg Pro Thr Phe
 470 475 480
 Glu Thr Leu Arg Trp Lys Leu Glu Asp Tyr Phe Glu Thr Asp Ser
 485 490 495
 20 Ser Tyr Ser Asp Ala Asn Asn Phe Ile Arg
 500 505

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 404 bases
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCGGCCGCAG AGAAAGCAGA GGATGGGGCT TAGCAGCTGG CAGAGCCAGG 50
 AGCGGGGAGG TAGCAGAAAG ACCACAAGTA CAAAGAAGTC CTGAAACTTT 100
 30 GGTTTTGCTG CTGCAGCCCA TTGAGAGTGA CGACATGGAG CACAAGACCC 150
 TGAAGATCAC CGACTTTGGC CTGGCCCGAG AGTGGCACAA AACCACACAA 200
 ATGAGTGCCG CNGGCACCTA CNCCTGGATG GCTCCTGAGG TTATCAAGGC 250
 CTCCACCTTC TCTAAGGGCA GTGACGTCTG GAGTTTGGG GTGCTGCTGT 300

GGGAACTGCT GACCGGGGAG NTGCCATACC GTGGCATTGA CTGCCTTGCT 350

GTGGCCTATG GCGTAGCTGT TAACAAGCTC AACTGCCAT CCATCCACCT 400

GGCC 404

(2) INFORMATION FOR SEQ ID NO:22:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3120 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATGAGAGCGT TGGCGCGCGA CGGCGGCCAG CTGCCGCTGC TCGTTGTTTT 50

TTCTGCAATG ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA 100

AGTGTGTTTT AATCAATCAT AAGAACAATG ATTCATCAGT GGGGAAGTCA 150

TCATCATATC CCATGGTATC AGAATCCCCG GAAGACCTCG GGTGTGCGTT 200

15 GAGACCCCAG AGCTCAGGGA CAGTGACGA AGCTGCCGCT GTGGAAGTGG 250

ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300

ATTCCTGTC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA 350

TTTTGATTTA CAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA 400

CAGAAACCCA AGCTGGAGAA TACCTACTTT TTATTCAGAG TGAAGCTACC 450

20 AATTACACAA TATTGTTTAC AGTGAGTATA AGAAATACCC TGCTTTACAC 500

ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC GCCCTGGTCT 550

GCATATCTGA GAGCGTTCCA GAGCGGATCC TGAATGGGT GCTTTGCGAT 600

TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TAAAAAAGGA 650

GGAAAAAGTG CTTTCATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA 700

GAAATGAACT GGGCAGGGAA TGCACCAGGC TGTTCACAAT AGATCTAAAT 750

CAAACCTCTC AGACCACATT GCCACAATTA TTTCTTAAAG TAGGGGAACC 800

5 CTTATGGATA AGGTGCAAAG CTGTTTCATGT GAACCATGGA TTCGGGCTCA 850

CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900

AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT 950

ATCATCAGTG GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA 1000

AGCATCCCAG TCAATCAGCT TTGGTTACCA TCGTAGAAAA GGGATTTATA 1050

10 AATGCTACCA ATTCAAGTGA AGATTATGAA ATTGACCAAT ATGAAGAGTT 1100

TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA TGTACGTGGA 1150

CCTTCTCTCG AAAATCATTT CTTGTGAGC AAAAGGGTCT TGATAACGGA 1200

TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT 1250

ATTCCATGCA GAAAATGATG ATGCCCAATT TACCAAATG TTCACGCTGT 1300

15 ATATAAGAAG GAAACCTCAA GTCCTCGCAG AAGCTTCGGC AAGTCAGGCG 1350

TCCTGTTTCT CGGATGGATA CCCATTACCA TCTTGGACCT GGAAGAAGTG 1400

TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA GGAGTCTGGA 1450

ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500

CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA 1550

CAATTCCCTT GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC 1600

CCTTCCCTTT CATCCAAGAC AACATCTCAT TCTATGCAAC AATTGGTGTT 1650

TGTCTCCTCT TCATTGTCGT TTTAACCCTG CTAATTTGTC ACAAGTACAA 1700

AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG GTGACCGGAT 1750

5 CCTCAGATTA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800

GTCAAATGGG AGTTTCCAAG AGAAAATTTA GAGTTTGGGA AGGTACTAGG 1850

ATCAGGTGCT TTTGGAAAAG TGATGAACGC AACAGCTTAT GGAATTAGCA 1900

AAACAGGAGT CTCAATCCAG GTTACCGTCA AAATGCTGAA AGAAAAAGCA 1950

GACAGCTCTG AAAGAGAGGC ACTCATGTCA GAACTCAAGA TGATGACCCA 2000

10 GCTGGGAAGC CACGAGAATA TTGTGAACCT GCTGGGGGCG TGCACACTGT 2050

CAGGACCAAT TTAATTGATT TTTGAATACT GTTGCTATGG TGATCTTCTC 2100

AACTATCTAA GAAGTAAAAG AGAAAAATTT CACAGGACTT GGACAGAGAT 2150

TTTCAAGGAA CACAATTTCA GTTTTTACCC CACTTTCCAA TCACATCCAA 2200

ATTCCAGCAT GCCTGGTTCA AGAGAAGTTC AGATACACCC GGACTCGGAT 2250

15 CAAATCTCAG GGCTTCATGG GAATTCATTT CACTCTGAAG ATGAAATTGA 2300

ATATGAAAAC CAAAAAAGGC TGGAAGAAGA GGAGGACTTG AATGTGCTTA 2350

CATTTGAAGA TCTTCTTTGC TTTGCATATC AAGTTGCCAA AGGAATGGAA 2400

TTTCTGGAAT TTAAGTCGTG TGTTACACAGA GACCTGGCCG CCAGGAACGT 2450

GCTTGTCACC CACGGGAAAAG TGGTGAAGAT ATGTGACTTT GGATTGGCTC 2500

GAGATATCAT GAGTGATTCC AACTATGTTG TCAGGGGCAA TGCCCGTCTG 2550
CCTGTAAAAT GGATGGCCCC CGAAAGCCTG TTTGAAGGCA TCTACACCAT 2600
TAAGAGTGAT GTCTGGTCAT ATGGAATATT ACTGTGGGAA ATCTTCTCAC 2650
TTGGTGTGAA TCCTTACCCT GGCATTCCGG TTGATGCTAA CTTCTACAAA 2700
5 CTGATTCAAA ATGGATTAA AATGGATCAG CCATTTTATG CTACAGAAGA 2750
AATATACATT ATAATGCAAT CCTGCTGGGC TTTTGAAGTCA AGGAAACGGC 2800
CATCCTTCCC TAATTTGACT TCGTTTTTAG GATGTCAGCT GGCAGATGCA 2850
GAAGAAGCGA TGTATCAGAA TGTGGATGGC CGTGTTCGG AATGTCCTCA 2900
CACCTACCAA AACAGGCGAC CTTTCAGCAG AGAGATGGAT TTGGGGCTAC 2950
10 TCTCTCCGCA GGCTCAGGTC GAAGATTCGT AGAGGAACAA TTTAGTTTTA 3000
AGGACTTCAT CCCTCCACCT ATCCCTAACA GGCTGTAGAT TACCAAAACA 3050
AGGTTAATTT CATCACTAAA AGAAAATCTA TTATCAACTG CTGCTTCACC 3100
AGACTTTTCT CTAGAGAGCG 3120

(2) INFORMATION FOR SEQ ID NO:23:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3969 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCGGCGTCCA CCCGCCAGG GAGAGTCAGA CCTGGGGGGG CGAGGGCCCC 50
CCAAACTCAG TTCGGATCCT ACCCGAGTGA GGCGGCGCCA TGGAGCTCCG 100

GGTGCTGCTC TGCTGGGCTT CGTTGGCCGC AGCTTTGGAA GAGACCCTGC 150

TGAACACAAA ATTGGAAACT GCTGATCTGA AGTGGGTGAC ATTCCCTCAG 200

GTGGACGGGC AGTGGGAGGA ACTGAGCGGC CTGGATGAGG AACAGCACAG 250

CGTGCGCACC TACGAAGTGT GTGACGTGCA GCGTGCCCCG GGCCAGGCCC 300

5 ACTGGCTTCG CACAGGTTGG GTCCACCGGC GGGGCGCCGT CCACGTGTAC 350

GCCACGCTGC GCTTCACCAT GCTCGAGTGC CTGTCCCTGC CTCGGGCTGG 400

GCGCTCCTGC AAGGAGACCT TCACCGTCTT CTACTATGAG AGCGATGCGG 450

ACACGGCCAC GGCCCTCACG CCAGCCTGGA TGGAGAACCC CTACATCAAG 500

GTGGACACGG TGGCCGCGGA GCATCTCACC CGGAAGCGCC CTGGGGCCGA 550

10 GGCCACCGGG AAGGTGAATG TCAAGACGCT GCGTCTGGGA CCGCTCAGCA 600

AGGCTGGCTT CTACCTGGCC TTCCAGGACC AGGGTGCCCTG CATGGCCCTG 650

CTATCCCTGC ACCTCTTCTA CAAAAGTGC GCCCAGCTGA CTGTGAACCT 700

GACTCGATTC CCGGAGACTG TGCCTCGGGA GCTGGTTGTG CCCGTGGCCG 750

GTAGCTGCGT GGTGGATGCC GTCCCCGCC CTGGCCCCAG CCCCAGCCTC 800

15 TACTGCCGTG AGGATGGCCA GTGGGCCGAA CAGCCGGTCA CGGGCTGCAG 850

CTGTGCTCCG GGGTTCGAGG CAGCTGAGGG GAACACCAAG TGCCGAGCCT 900

GTGCCCAGGG CACCTTCAAG CCCCTGTCAG GAGAAGGGTC CTGCCAGCCA 950

TGCCAGCCA ATAGCCACTC TAACACCATT GGATCAGCCG TCTGCCAGTG 1000

CCGCGTCGGG TACTTCCGGG CACGCACAGA CCCCCGGGGT GCACCCTGCA 1050

WO 95/27061

CCACCCCTCC TTCGGCTCCG CGGAGCGTGG TTTCCCGCCT GAACGGCTCC 1100
 TCCCTGCACC TGGAATGGAG TGCCCCCTG GAGTCTGGTG GCCGAGAGGA 1150
 CCTCACCTAC GCCCTCCGCT GCCGGGAGTG CCGACCCGGA GGCTCCTGTG 1200
 CGCCCTGCGG GGGAGACCTG ACTTTTGACC CCGGCCCCCG GGACCTGGTG 1250
 5 GAGCCCTGGG TGGTGGTTCG AGGGCTACGT CCTGACTTCA CCTATACCTT 1300
 TGAGGTCACT GCATTGAACG GGGTATCCTC CTTAGCCACG GGGCCCGTCC 1350
 CATTTGAGCC TGTCAATGTC ACCACTGACC GAGAGGTACC TCCTGCAGTG 1400
 TCTGACATCC GGGTGACGCG GTCCTCACCC AGCAGCTTGA GCCTGGCCTG 1450
 GGCTGTTCCC CGGGCACCCA GTGGGGCTGT GCTGGACTAC GAGGTCAAAT 1500
 10 ACCATGAGAA GGGCGCCGAG GGTCCCAGCA GCGTGCGGTT CCTGAAGACG 1550
 TCAGAAAACC GGGCAGAGCT GCGGGGGCTG AAGCGGGGAG CCAGCTACCT 1600
 GGTGCAGGTA CGGGCGCGCT CTGAGGCCGG CTACGGGCCC TTCGGCCAGG 1650
 AACATCACAG CCAGACCCAA CTGGATGAGA GCGAGGGCTG GCGGGAGCAG 1700
 CTGGCCCTGA TTGCGGGCAC GGCAGTCGTG GGTGTGGTCC TGGTCCTGGT 1750
 15 GGTCATTGTG GTCGCAGTTC TCTGCCTCAG GAAGCAGAGC AATGGGAGAG 1800
 AAGCAGAATA TTCGGACAAA CACGGACAGT ATCTCATCGG ACATGGTACT 1850
 AAGGTCTACA TCGACCCCTT CACTTATGAA GACCCTAATG AGGCTGTGAG 1900
 GGAATTTGCA AAAGAGATCG ATGTCTCCTA CGTCAAGATT GAAGAGGTGA 1950
 TTGGTGCAGG TGAGTTTGGC GAGGTGTGCC GGGGGCGGCT CAAGGCCCCA 2000

GGGAAGAAGG AGAGCTGTGT GGCAATCAAG ACCCTGAAGG GTGGCTACAC 2050

GGAGCGGCAG CGGCGTGAGT TTCTGAGCGA GGCCTCCATC ATGGGCCAGT 2100

TCGAGCACCC CAATATCATC CGCCTGGAGG GCGTGGTCAC CAACAGCATG 2150

CCCGTCATGA TTCTCACAGA GTTCATGGAG AACGGCGCCC TGGACTCCTT 2200

5 CCTGCGGCTA AACGACGGAC AGTTCACAGT CATCCAGCTC GTGGGCATGC 2250

TGCGGGGCAT CGCCTCGGGC ATGCGGTACC TTGCCGAGAT GAGCTACGTC 2300

CACCGAGACC TGGCTGCTCG CAACATCCTA GTCAACAGCA ACCTCGTCTG 2350

CAAAGTGTCT GACTTTGGCC TTTCCCGATT CCTGGAGGAG AACTCTTCCG 2400

ATCCCACCTA CACGAGCTCC CTGGGAGGAA AGATTCCCAT CCGATGGACT 2450

10 GCCCCGGAGG CCATTGCCTT CCGGAAGTTC ACTTCCGCCA GTGATGCCTG 2500

GAGTTACGGG ATTGTGATGT GGGAGGTGAT GTCATTTGGG GAGAGGCCGT 2550

ACTGGGACAT GAGCAATCAG GACGTGATCA ATGCCATTGA ACAGGACTAC 2600

CGGCTGCCCC CGCCCCCAGA CTGTCCCACC TCCCTCCACC AGCTCATGCT 2650

GGACTGTTGG CAGAAAGACC GGAATGCCCC GCCCCGCTTC CCCCAGGTGG 2700

15 TCAGCGCCCT GGACAAGATG ATCCGGAACC CCGCCAGCCT CAAAATCGTG 2750

GCCCCGGAGA ATGGCGGGGC CTCACACCCT CTCCTGGACC AGCGGCAGCC 2800

TCACTACTCA GCTTTTGGCT CTGTGGGCGA GTGGCTTCGG GCCATCAAAA 2850

TGGGAAGATA CGAAGAAAGT TTCGCAGCCG CTGGCTTTGG CTCCTTCGAG 2900

CTGGTCAGCC AGATCTCTGC TGAGGACCTG CTCCGAATCG GAGTCACTCT 2950

GGCGGGACAC CAGAAGAAAA TCTTGGCCAG TGTCCAGCAC ATGAAGTCCC 3000

AGGCCAAGCC GGGAACCCCG GGTGGGACAG GAGGACCGGC CCCGCAGTAC 3050

TGACCTGCAG GAACTCCCCA CCCAGGGAC ACCGCCTCCC CATTTTCCGG 3100

GGCAGAGTGG GGA CTCACAG AGGCCCCCAG CCCTGTGCCC CGCTGGATTG 3150

5 CACTTTGAGC CCGTGGGGTG AGGAGTTGGC AATTGGAGA GACAGGATTT 3200

GGGGGTTCTG CCATAATAGG AGGGGAAAAT CACCCCCCAG CCACCTCGGG 3250

GAACTCCAGA CCAAGGGTGA GGGCGCCTTT CCCTCAGGAC TGGGTGTGAC 3300

CAGAGGAAAA GGAAGTGCCC AACATCTCCC AGCCTCCCCA GGTGCCCCCC 3350

TCACCTTGAT GGGTGCGTTC CCGCAGACCA AAGAGAGTGT GACTCCCTTG 3400

10 CCAGCTCCAG AGTGGGGGGG CTGTCCCAGG GGGCAAGAAG GGGTGTGAGG 3450

GCCCAGTGAC AAAATCATTG GGGTTGTAG TCCCAACTTG CTGCTGTCAC 3500

CACCAAATC AATCATTTTT TTCCCTTGTA AATGCCCCCTC CCCCAGCTGC 3550

TGCCTTCATA TTGAAGGTTT TTGAGTTTGT TTTTGGTCT TAATTTTCT 3600

CCCCGTTCCC TTTTGTTC TTCGTTTTGT TTTCTACCG TCCTTGTCAT 3650

15 AACTTTGTGT TGGAGGGAAC CTGTTTCACT ATGGCCTCCT TGCCCCAAGT 3700

TGAAACAGGG GCCCATCATC ATGTCTGTTT CCAGAACAGT GCCTTGGTCA 3750

TCCCACATCC CCGGACCCCG CCTGGGACCC CCAAGCTGTG TCCTATGAAG 3800

GGGTGTGGGG TGAGGTAGTG AAAAGGGCGG TAGTTGGTGG TGGAACCCAG 3850

AAACGGACGC CGGTGCTTGG AGGGGTCTT AAATTATATT TAAAAAAGTA 3900

ACTTTTGTGTA TAAATAAAAG AAAATGGGAC GTGTCCCAGC TCCAGGGGTA 3950

AAAAAAAAAAAA 3969

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1276 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

10	Met	Glu	Leu	Arg	Val	Leu	Leu	Cys	Trp	Ala	Ser	Leu	Ala	Ala	Ala	1	5	10	15
	Leu	Glu	Glu	Thr	Leu	Leu	Asn	Thr	Lys	Leu	Glu	Thr	Ala	Asp	Leu	20	25	30	
	Lys	Trp	Val	Thr	Phe	Pro	Gln	Val	Asp	Gly	Gln	Trp	Glu	Glu	Leu	35	40	45	
15	Ser	Gly	Leu	Asp	Glu	Glu	Gln	His	Ser	Val	Arg	Thr	Tyr	Glu	Val	50	55	60	
	Cys	Asp	Val	Gln	Arg	Ala	Pro	Gly	Gln	Ala	His	Trp	Leu	Arg	Thr	65	70	75	
20	Gly	Trp	Val	Pro	Arg	Arg	Gly	Ala	Val	His	Val	Tyr	Ala	Thr	Leu	80	85	90	
	Arg	Phe	Thr	Met	Leu	Glu	Cys	Leu	Ser	Leu	Pro	Arg	Ala	Gly	Arg	95	100	105	
	Ser	Cys	Lys	Glu	Thr	Phe	Thr	Val	Phe	Tyr	Tyr	Glu	Ser	Asp	Ala	110	115	120	
25	Asp	Thr	Ala	Thr	Ala	Leu	Thr	Pro	Ala	Trp	Met	Glu	Asn	Pro	Tyr	125	130	135	
	Ile	Lys	Val	Asp	Thr	Val	Ala	Ala	Glu	His	Leu	Thr	Arg	Lys	Arg	140	145	150	
30	Pro	Gly	Ala	Glu	Ala	Thr	Gly	Lys	Val	Asn	Val	Lys	Thr	Leu	Arg	155	160	165	
	Leu	Gly	Pro	Leu	Ser	Lys	Ala	Gly	Phe	Tyr	Leu	Ala	Phe	Gln	Asp	170	175	180	
	Gln	Gly	Ala	Cys	Met	Ala	Leu	Leu	Ser	Leu	His	Leu	Phe	Tyr	Lys	185	190	195	
35	Lys	Cys	Ala	Gln	Leu	Thr	Val	Asn	Leu	Thr	Arg	Phe	Pro	Glu	Thr	200	205	210	

	Val	Pro	Arg	Glu	Leu	Val	Val	Pro	Val	Ala	Gly	Ser	Cys	Val	Val	215	220	225
	Asp	Ala	Val	Pro	Ala	Pro	Gly	Pro	Ser	Pro	Ser	Leu	Tyr	Cys	Arg	230	235	240
5	Glu	Asp	Gly	Gln	Trp	Ala	Glu	Gln	Pro	Val	Thr	Gly	Cys	Ser	Cys	245	250	255
	Ala	Pro	Gly	Phe	Glu	Ala	Ala	Glu	Gly	Asn	Thr	Lys	Cys	Arg	Ala	260	265	270
10	Cys	Ala	Gln	Gly	Thr	Phe	Lys	Pro	Leu	Ser	Gly	Glu	Gly	Ser	Cys	275	280	285
	Gln	Pro	Cys	Pro	Ala	Asn	Ser	His	Ser	Asn	Thr	Ile	Gly	Ser	Ala	290	295	300
	Val	Cys	Gln	Cys	Arg	Val	Gly	Tyr	Phe	Arg	Ala	Arg	Thr	Asp	Pro	305	310	315
15	Arg	Gly	Ala	Pro	Cys	Thr	Thr	Pro	Pro	Ser	Ala	Pro	Arg	Ser	Val	320	325	330
	Val	Ser	Arg	Leu	Asn	Gly	Ser	Ser	Leu	His	Leu	Glu	Trp	Ser	Ala	335	340	345
20	Pro	Leu	Glu	Ser	Gly	Gly	Arg	Glu	Asp	Leu	Thr	Tyr	Ala	Leu	Arg	350	355	360
	Cys	Arg	Glu	Cys	Arg	Pro	Gly	Gly	Ser	Cys	Ala	Pro	Cys	Gly	Gly	365	370	375
	Asp	Leu	Thr	Phe	Asp	Pro	Gly	Pro	Arg	Asp	Leu	Val	Glu	Pro	Trp	380	385	390
25	Val	Val	Val	Arg	Gly	Leu	Arg	Pro	Asp	Phe	Thr	Tyr	Thr	Phe	Glu	395	400	405
	Val	Thr	Ala	Leu	Asn	Gly	Val	Ser	Ser	Leu	Ala	Thr	Gly	Pro	Val	410	415	420
30	Pro	Phe	Glu	Pro	Val	Asn	Val	Thr	Thr	Asp	Arg	Glu	Val	Pro	Pro	425	430	435
	Ala	Val	Ser	Asp	Ile	Arg	Val	Thr	Arg	Ser	Ser	Pro	Ser	Ser	Leu	440	445	450
	Ser	Leu	Ala	Trp	Ala	Val	Pro	Arg	Ala	Pro	Ser	Gly	Ala	Val	Leu	455	460	465
35	Asp	Tyr	Glu	Val	Lys	Tyr	His	Glu	Lys	Gly	Ala	Glu	Gly	Pro	Ser	470	475	480
	Ser	Val	Arg	Phe	Leu	Lys	Thr	Ser	Glu	Asn	Arg	Ala	Glu	Leu	Arg	485	490	495

	Gly Leu Lys Arg Gly Ala Ser Tyr Leu Val Gln Val Arg Ala Arg	500	505	510
	Ser Glu Ala Gly Tyr Gly Pro Phe Gly Gln Glu His His Ser Gln	515	520	525
5	Thr Gln Leu Asp Glu Ser Glu Gly Trp Arg Glu Gln Leu Ala Leu	530	535	540
	Ile Ala Gly Thr Ala Val Val Gly Val Val Leu Val Leu Val Val	545	550	555
10	Ile Val Val Ala Val Leu Cys Leu Arg Lys Gln Ser Asn Gly Arg	560	565	570
	Glu Ala Glu Tyr Ser Asp Lys His Gly Gln Tyr Leu Ile Gly His	575	580	585
	Gly Thr Lys Val Tyr Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn	590	595	600
15	Glu Ala Val Arg Glu Phe Ala Lys Glu Ile Asp Val Ser Tyr Val	605	610	615
	Lys Ile Glu Glu Val Ile Gly Ala Gly Glu Phe Gly Glu Val Cys	620	625	630
20	Arg Gly Arg Leu Lys Ala Pro Gly Lys Lys Glu Ser Cys Val Ala	635	640	645
	Ile Lys Thr Leu Lys Gly Gly Tyr Thr Glu Arg Gln Arg Arg Glu	650	655	660
	Phe Leu Ser Glu Ala Ser Ile Met Gly Gln Phe Glu His Pro Asn	665	670	675
25	Ile Ile Arg Leu Glu Gly Val Val Thr Asn Ser Met Pro Val Met	680	685	690
	Ile Leu Thr Glu Phe Met Glu Asn Gly Ala Leu Asp Ser Phe Leu	695	700	705
30	Arg Leu Asn Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met	710	715	720
	Leu Arg Gly Ile Ala Ser Gly Met Arg Tyr Leu Ala Glu Met Ser	725	730	735
	Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser	740	745	750
35	Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Phe Leu	755	760	765
	Glu Glu Asn Ser Ser Asp Pro Thr Tyr Thr Ser Ser Leu Gly Gly	770	775	780

	Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Phe Arg	785	790	795
	Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile Val Met	800	805	810
5	Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser	815	820	825
	Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro	830	835	840
10	Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp	845	850	855
	Cys Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val	860	865	870
	Val Ser Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys	875	880	885
15	Ile Val Ala Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp	890	895	900
	Gln Arg Gln Pro His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp	905	910	915
20	Leu Arg Ala Ile Lys Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala	920	925	930
	Ala Gly Phe Gly Ser Phe Glu Leu Val Ser Gln Ile Ser Ala Glu	935	940	945
	Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys Lys	950	955	960
25	Ile Leu Ala Ser Val Gln His Met Lys Ser Gln Ala Lys Pro Gly	965	970	975
	Thr Pro Gly Gly Thr Gly Gly Pro Ala Pro Gln Tyr Pro Ala Gly	980	985	990
30	Thr Pro His Pro Arg Asp Thr Ala Ser Pro Phe Ser Gly Ala Glu	995	1000	1005
	Trp Gly Leu Thr Glu Ala Pro Ser Pro Val Pro Arg Trp Ile Ala	1010	1015	1020
	Leu Ala Arg Gly Val Arg Ser Trp Gln Phe Gly Glu Thr Gly Phe	1025	1030	1035
35	Gly Gly Ser Ala Ile Ile Gly Gly Glu Asn His Pro Pro Ala Thr	1040	1045	1050
	Ser Gly Asn Ser Arg Pro Arg Val Arg Ala Pro Phe Pro Gln Asp	1055	1060	1065

Trp Val Pro Glu Glu Lys Glu Val Pro Asn Ile Ser Gln Pro Pro
 1070 1075 1080
 Gln Val Pro Pro Ser Pro Trp Val Arg Ser Arg Arg Pro Lys Arg
 1085 1090 1095
 5 Val Leu Pro Cys Gln Leu Gln Ser Gly Gly Ala Val Pro Gly Gly
 1100 1105 1110
 Lys Lys Gly Cys Gln Gly Pro Val Thr Lys Ser Leu Gly Phe Val
 1115 1120 1125
 Val Pro Thr Cys Cys Cys His His Gln Thr Gln Ser Phe Phe Ser
 10 1130 1135 1140
 Leu Val Asn Ala Pro Pro Pro Ala Ala Ala Phe Ile Leu Lys Val
 1145 1150 1155
 Phe Glu Phe Cys Phe Trp Ser Phe Phe Ser Pro Phe Pro Phe Cys
 1160 1165 1170
 15 Phe Phe Val Leu Phe Phe Tyr Arg Pro Cys His Asn Phe Val Leu
 1175 1180 1185
 Glu Gly Thr Cys Phe Thr Met Ala Ser Phe Ala Gln Val Glu Thr
 1190 1195 1200
 Gly Ala His His His Val Cys Phe Gln Asn Ser Ala Leu Val Ile
 20 1205 1210 1215
 Pro His Pro Arg Thr Pro Pro Gly Thr Pro Lys Leu Cys Pro Met
 1220 1225 1230
 Lys Gly Cys Gly Val Arg Lys Gly Arg Leu Val Val Glu Pro Arg
 1235 1240 1245
 25 Asn Gly Arg Arg Cys Leu Glu Gly Phe Leu Asn Tyr Ile Lys Ser
 1250 1255 1260
 Asn Phe Leu Tyr Lys Lys Lys Met Gly Arg Val Pro Ala Pro Gly
 1265 1270 1275
 Val
 30 1276

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser
 1 5 10 15

PCT/US95/04228

(2) INFORMATION FOR SEQ ID NO:26:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

(2) INFORMATION FOR SEQ ID NO:27:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

35 (2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Asn Cys Leu Val Gly Lys Asn Tyr Thr Ile Lys Ile Ala Asp Phe
 1 5 10 15
 Gly Met Ser Arg Asn Leu Tyr Ser Gly Asp Tyr Tyr
 5 20 25 27

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Thr Arg Asn Ile Leu Val Glu Asn Glu Asn Arg Val Lys Ile Gly
 1 5 10 15
 Asp Phe Gly Leu Thr Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr
 15 20 25 30
 Lys Val Lys Glu Pro Gly Glu Ser Pro Ile Phe Trp Tyr Ala Pro
 35 40 45
 Glu Ser Leu Thr Glu Ser Leu Phe Ser Val Ala Ser Asp
 50 55 58

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ala Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser
 1 5 10 15
 Asp Phe Gly Met Ser Arg Val Leu Glu Asp Asp Pro Glu Ala Ala
 20 25 30
 Tyr Thr Thr Arg Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala Pro
 30 35 40 45
 Glu Ala Ile Ala Tyr Arg Lys Phe Thr Ser Ala Ser Asp
 50 55 58

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4425 bases
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCGGGTCGGA CCCACGCGCA GCGGCCGGAG ATGCAGCGGG GCGCCGCGCT 50

GTGCCTGCGA CTGTGGCTCT GCCTGGGACT CCTGGACGGC CTGGTGAGTG 100

GCTACTCCAT GACCCCCCG ACCTTGAACA TCACGGAGGA GTCACACGTC 150

5 ATCGACACCG GTGACAGCCT GTCCATCTCC TGCAGGGGAC AGCACCCT 200

CGAGTGGGCT TGGCCAGGAG CTCAGGAGGC GCCAGCCACC GGAGACAAGG 250

ACAGCGAGGA CACGGGGGTG GTGCGAGACT GCGAGGGCAC AGACGCCAGG 300

CCCTACTGCA AGGTGTTGCT GCTGCACGAG GTACATGCCA ACGACACAGG 350

CAGCTACGTC TGCTACTACA AGTACATCAA GGCACGCATC GAGGGCACCA 400

10 CGGCCGCCAG CTCCTACGTG TTCGTGAGAG ACTTTGAGCA GCCATTCATC 450

AACAAGCCTG ACACGCTCTT GGTCAACAGG AAGGACGCCA TGTGGGTGCC 500

CTGTCTGGTG TCCATCCCCG GCCTCAATGT CACGCTGCGC TCGCAAAGCT 550

CGGTGCTGTG GCCAGACGGG CAGGAGGTGG TGTGGGATGA CCGGCGGGGC 600

ATGCTCGTGT CCACGCCACT GCTGCACGAT GCCCTGTACC TGCAGTGCGA 650

15 GACCACCTGG GGAGACCAGG ACTTCCTTTC CAACCCCTTC CTGGTGACCA 700

TCACAGGCAA CGAGCTCTAT GACATCCAGC TGTTGCCCAG GAAGTCGCTG 750

GAGCTGCTGG TAGGGGAGAA GCTGGTCCTG AACTGCACCG TGTGGGCTGA 800

GTTTAACTCA GGTGTCACCT TTGACTGGGA CTACCCAGGG AAGCAGGCAG 850

AGCGGGGTAA GTGGGTGCCC GAGCGACGCT CCCAGCAGAC CCACACAGAA 900

CTCTCCAGCA TCCTGACCAT CCACAACGTC AGCCAGCACG ACCTGGGCTC 950

GTATGTGTGC AAGGCCAACA ACGGCATCCA GCGATTTCGG GAGAGCACCG 1000

AGGTCATTGT GCATGAAAAT CCCTTCATCA GCGTCGAGTG GCTCAAAGGA 1050

CCCATCCTGG AGGCCACGGC AGGAGACGAG CTGGTGAAGC TGCCCGTGAA 1100

5 GCTGGCAGCG TACCCCCCGC CCGAGTTCCA GTGGTACAAG GATGGAAAGG 1150

CACTGTCCGG GCGCCACAGT CCACATGCCC TGGTGCTCAA GGAGGTGACA 1200

GAGGCCAGCA CAGGCACCTA CACCCTCGCC CTGTGGA ACT CCGCTGCTGG 1250

CCTGAGGCGC AACATCAGCC TGGAGCTGGT GGTGAATGTG CCCCCCAGA 1300

TACATGAGAA GGAGGCCTCC TCCCCCAGCA TCTACTCGCG TCACAGCCGC 1350

10 CAGGCCCTCA CCTGCACGGC CTACGGGGTG CCCCTGCCTC TCAGCATCCA 1400

GTGGCACTGG CGGCCCTGGA CACCCTGCAA GATGTTTGCC CAGCGTAGTC 1450

TCCGGCGGCG GCAGCAGCAA GACCTCATGC CACAGTGCCG TGA CTGGAGG 1500

GCGGTGACCA CGCAGGATGC CGTGAACCCC ATCGAGAGCC TGGACACCTG 1550

GACCGAGTTT GTGGAGGGAA AGAATAAGAC TGTGAGCAAG CTGGTGATCC 1600

15 AGAATGCCAA CGTGTCTGCC ATGTACAAGT GTGTGGTCTC CAACAAGGTG 1650

GGCCAGGATG AGCGGCTCAT C TACTTCTAT GTGACCACCA TCCCCGACGG 1700

CTTCACCATC GAATCCAAGC CATCCGAGGA GCTACTAGAG GGCCAGCCGG 1750

TGCTCCTGAG CTGCCAAGCC GACAGCTACA AGTACGAGCA TCTGCGCTGG 1800

TACCGCCTCA ACCTGTCCAC GCTGCACGAT GCGCACGGGA ACCCGCTTCT 1850

GCTCGACTGC AAGAACGTGC ATCTGTTTCGC CACCCCTCTG GCCGCCAGCC 1900

TGGAGGAGGT GGCACCTGGG GCGCGCCACG CCACGCTCAG CCTGAGTATC 1950

CCCCGCGTCG CGCCCCGAGCA CGAGGGCCAC TATGTGTGCG AAGTGCAAGA 2000

CCGGCGCAGC CATGACAAGC ACTGCCACAA GAAGTACCTG TCGGTGCAGG 2050

5 CCCTGGAAGC CCCTCGGCTC ACGCAGAACT TGACCGACCT CCTGGTGAAC 2100

GTGAGCGACT CGCTGGAGAT GCAGTGCTTG GTGGCCGGAG CGCACGCGCC 2150

CAGCATCGTG TGGTACAAAG ACGAGAGGCT GCTGGAGGAA AAGTCTGGAG 2200

TCGACTTGGC GGA CTCCAAC CAGAAGCTGA GCATCCAGCG CGTGCGCGAG 2250

GAGGATGCGG GACGCTATCT GTGCAGCGTG TGCAACGCCA AGGGCTGCGT 2300

10 CAACTCCTCC GCCAGCGTGG CCGTGGAAGG CTCCGAGGAT AAGGGCAGCA 2350

TGGAGATCGT GATCCTTGTC GGTACCGGCG TCATCGCTGT CTTCTTCTGG 2400

GTCCTCCTCC TCCTCATCTT CTGTAACATG AGGAGGCCGG CCCACGCAGA 2450

CATCAAGACG GGCTACCTGT CCATCATCAT GGACCCCGGG GAGGTGCCTC 2500

TGGAGGAGCA ATGCGAATAC CTGTCCTACG ATGCCAGCCA GTGGGAATTC 2550

15 CCCCCGAGAGC GGCTGCACCT GGGGAGAGTG CTCGGCTACG GCGCCTTCGG 2600

GAAGGTGGTG GAAGCCTCCG CTTTCGGCAT CCACAAGGGC AGCAGCTGTG 2650

ACACCGTGGC CGTGAAAATG CTGAAAGAGG GCGCCACGGC CAGCGAGCAC 2700

CGCGCGCTGA TGTCGGAGCT CAAGATCCTC ATTCACATCG GCAACCACCT 2750

CAACGTGGTC AACCTCCTCG GGGCGTGCAC CAAGCCGCAG GGCCCCCTCA 2800

TGGTGATCGT GGAGTTCTGC AAGTACGGCA ACCTCTCAA CTTCTGCGC 2850

GCCAAGCGGG ACGCCTTCAG CCCCTGCGCG GAGAAGTCTC CCGAGCAGCG 2900

CGGACGCTTC CGCGCCATGG TGGAGCTCGC CAGGCTGGAT CGGAGGCGGC 2950

CGGGGAGCAG CGACAGGGTC CTCTTCGCGC GGTTCTCGAA GACCGAGGGC 3000

5 GGAGCGAGGC GGGCTTCTCC AGACCAAGAA GCTGAGGACC TGTGGCTGAG 3050

CCCGCTGACC ATGGAAGATC TTGTCTGCTA CAGCTTCCAG GTGGCCAGAG 3100

GGATGGAGTT CCTGGCTTCC CGAAAGTGCA TCCACAGAGA CCTGGCTGCT 3150

CGGAACATTC TGCTGTCGGA AAGCGACGTG GTGAAGATCT GTGACTTTGG 3200

CCTTGCCCGG GACATCTACA AAGACCCTGA CTACGTCCGC AAGGGCAGTG 3250

10 CCCGGCTGCC CCTGAAGTGG ATGGCCCCTG AAAGCATCTT CGACAAGGTG 3300

TACACCACGC AGAGTGACGT GTGGTCCTTT GGGGTGCTTC TCTGGGAGAT 3350

CTTCTCTCTG GGGGCCTCCC CGTACCCTGG GGTGCAGATC AATGAGGAGT 3400

TCTGCCAGCG GCTGAGAGAC GGCACAAGGA TGAGGGCCCC GGAGCTGGCC 3450

ACTCCCGCCA TACGCCGCAT CATGCTGAAC TGCTGGTCCG GAGACCCCAA 3500

15 GGCGAGACCT GCATTCTCGG AGCTGGTGGA GATCCTGGGG GACCTGCTCC 3550

AGGGCAGGGG CCTGCAAGAG GAAGAGGAGG TCTGCATGGC CCCGCGCAGC 3600

TCTCAGAGCT CAGAAGAGGG CAGCTTCTCG CAGGTGTCCA CCATGGCCCT 3650

ACACATCGCC CAGGCTGACG CTGAGGACAG CCCGCCAAGC CTGCAGCGCC 3700

ACAGCCTGGC CGCCAGGTAT TACAACCTGGG TGTCTTTCC CGGGTGCCTG 3750

GCCAGAGGGG CTGAGACCCG TGGTTCTCTCC AGGATGAAGA CATTTGAGGA 3800

ATTCCCCATG ACCCCAACGA CCTACAAAGG CTCTGTGGAC AACCAGACAG 3850

ACAGTGGGAT GGTGCTGGCC TCGGAGGAGT TTGAGCAGAT AGAGAGCAGG 3900

CATAGACAAG AAAGCGGCTT CAGGTAGCTG AAGCAGAGAG AGAGAAGGCA 3950

5 GCATACGTCA GCATTTTCTT CTCTGCACTT ATAAGAAAGA TCAAAGACTT 4000

TAAGACTTTC GCTATTTCTT CTGCTATCTA CTACAAACTT CAAAGAGGAA 4050

CCAGGAGGCC AAGAGGAGCA TGAAAGTGGG CAAGGAGTGT GACCACTGAA 4100

GCACCACAGG GAGGGGTTAG GCCTCCGGAT GACTGCGGGC AGGCCTGGAT 4150

AATATCCAGC CTCCCACAAG AAGCTGGTGG AGCAGAGTGT TCCCTGACTC 4200

10 CTCCAAGGAA AGGGAGACGC CCTTTCATGG TCTGCTGAGT AACAGGTGCC 4250

TTCCCAGACA CTGGCGTTAC TGCTTGACCA AAGAGCCCTC AAGCGGCCCT 4300

TATGCCAGCG TGACAGAGGG CTCACCTCTT GCCTTCTAGG TCACTTCTCA 4350

CAATGTCCCT TCAGCACCTG ACCCTGTGCC CGCCAGTTAT TCCTTGGTAA 4400

TATGAGTAAT ACATCAAAGA GTAGT 4425

15 (2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 4425 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCCCAGCCT GGGTGC GCGT CGCCGGCCCTC TACGTCGCCC CGCGGCGCGA 50

CACGGACGCT GACACCGAGA CGGACCCTGA GGACCTGCCG GACCACTCAC 100

CGATGAGGTA CTGGGGGGGC TGGAACTTGT AGTGCCTCCT CAGTGTGCAG 150

TAGCTGTGGC CACTGTCGGA CAGGTAGAGG ACGTCCCCTG TCGTGGGGGA 200

GCTCACCCGA ACCGGTCCTC GAGTCCTCCG CGGTCCGTGG CCTCTGTTCC 250

5 TGTCGCTCCT GTGCCCCCAC CACGCTCTGA CGCTCCCGTG TCTGCGGTCC 300

GGGATGACGT TCCACAACGA CGACGTGCTC CATGTACGGT TGCTGTGTCC 350

GTCGATGCAG ACGATGATGT TCATGTAGTT CCGTGCGTAG CTCCCGTGGT 400

GCCGGCGGTC GAGGATGCAC AAGCACTCTC TGAAACTCGT CGGTAAGTAG 450

TTGTTCCGAC TGTGCGAGAA CCAGTTGTCC TTCCTGCGGT ACACCCACGG 500

10 GACAGACCAC AGGTAGGGGC CGGAGTTACA GTGCGACGCG AGCGTTTCGA 550

GCCACGACAC CGGTCTGCCC GTCCTCCACC ACACCCTACT GGCCGCCCCG 600

TACGAGCACA GGTGCGGTGA CGACGTGCTA CGGGACATGG ACGTCACGCT 650

CTGGTGGACC CCTCTGGTCC TGAAGGAAAG GTTGGGGAAG GACCACGTGT 700

AGTGTCCGTT GCTCGAGATA CTGTAGGTCG ACAACGGGTC CTTCAGCGAC 750

15 CTCGACGACC ATCCCCTCTT CGACCAGGAC TTGACGTGGC ACACCCGACT 800

CAAATTGAGT CCACAGTGGA AACTGACCCT GATGGGTCCC TTCGTCCGTC 850

TCGCCCCATT CACCCACGGG CTCGCTGCGA GGGTCGTCTG GGTGTGTCTT 900

GAGAGGTCGT AGGACTGGTA GGTGTTGCAG TCGGTCTGTC TGGACCCGAG 950

CATACACACG TTCCGGTTGT TGCCGTAGGT CGCTAAAGCC CTCTCGTGGC 1000

TCCAGTAACA CGTACTTTTA GGGAAAGTAGT CGCAGCTCAC CGAGTTTCCT 1050

GGGTAGGACC TCCGGTGCCG TCCTCTGCTC GACCACTTCG ACGGGCACTT 1100

CGACCGTCGC ATGGGGGGCG GGCTCAAGGT CACCATGTTC CTACCTTTCC 1150

GTGACAGGCC CGCGGTGTCA GGTGTACGGG ACCACGAGTT CCTCCACTGT 1200

5 CTCCGGTCGT GTCCGTGGAT GTGGGAGCGG GACACCTTGA GGCGACGACC 1250

GGACTCCGCG TTGTAGTCGG ACCTCGACCA CCACTTACAC GGGGGGGTCT 1300

ATGTACTCTT CCTCCGGAGG AGGGGGTCTG AGATGAGCGC AGTGTCGGCG 1350

GTCCGGGAGT GGACGTGCCG GATGCCCCAC GGGGACGGAG AGTCGTAGGT 1400

CACCGTGACC GCCGGGACCT GTGGGACGTT CTACAAACGG GTCGCATCAG 1450

10 AGGCCGCCGC CGTCGTCGTT CTGGAGTACG GTGTCACGGC ACTGACCTCC 1500

CGCCACTGGT GCGTCCTACG GCACTTGGGG TAGCTCTCGG ACCTGTGGAC 1550

CTGGCTCAAA CACCTCCCTT TCTTATTCTG ACACTCGTTC GACCACTAGG 1600

TCTTACGGTT GCACAGACGG TACATGTTCA CACACCAGAG GTTGTTCCAC 1650

CCGGTCCTAC TCGCCGAGTA GATGAAGATA CACTGGTGGT AGGGGCTGCC 1700

15 GAAGTGGTAG CTTAGGTTCTG GTAGGCTCCT CGATGATCTC CCGGTCGGCC 1750

ACGAGGACTC GACGGTTCGG CTGTGATGT TCATGCTCGT AGACGCGACC 1800

ATGGCGGAGT TGGACAGGTG CGACGTGCTA CGCGTGCCCT TGGGCGAAGA 1850

CGAGCTGACG TTCTTGCACG TAGACAAGCG GTGGGGAGAC CGGCGGTCCG 1900

ACCTCCTCCA CCGTGGACCC CGCGCGGTGC GGTGCGAGTC GGACTCATAG 1950

GGGGCGCAGC GCGGGCTCGT GCTCCCGGTG ATACACACGC TTCACGTTCT 2000

GGCCGCGTCG GTACTGTTTCG TGACGGTGTT CTTCATGGAC AGCCACGTCC 2050

GGGACCTTCG GGGAGCCGAG TGCGTCTTGA ACTGGCTGGA GGACCACTTG 2100

CACTCGCTGA GCGACCTCTA CGTCACGAAC CACCGGCCTC GCGTGCGCGG 2150

5 GTCGTAGCAC ACCATGTTTC TGCTCTCCGA CGACCTCCTT TTCAGACCTC 2200

AGCTGAACCG CCTGAGGTTG GTCTTCGACT CGTAGGTCGC GCACGCGCTC 2250

CTCCTACGCC CTGCGATAGA CACGTGCGAC ACGTTGCGGT TCCCGACGCA 2300

GTTGAGGAGG CGGTGCGACC GGCACCTTCC GAGGCTCCTA TTCCCGTCGT 2350

ACCTCTAGCA CTAGGAACAG CCATGGCCGC AGTAGCGACA GAAGAAGACC 2400

10 CAGGAGGAGG AGGAGTAGAA GACATTGTAC TCCTCCGGCC GGGTGCGTCT 2450

GTAGTTCTGC CCGATGGACA GGTAGTAGTA CCTGGGGCCC CTCCACGGAG 2500

ACCTCCTCGT TACGCTTATG GACAGGATGC TACGGTCGGT CACCCTTAAG 2550

GGGGCTCTCG CCGACGTGGA CCCCTCTCAC GAGCCGATGC CGCGGAAGCC 2600

CTTCCACCAC CTTGCGAGGC GAAAGCCGTA GGTGTTCCCG TCGTCGACAC 2650

15 TGTGGCACCG GCACTTTTAC GACTTTCTCC CGCGGTGCCG GTCGCTCGTG 2700

GCGCGCGACT ACAGCCTCGA GTTCTAGGAG TAAGTGTAGC CGTTGGTGGA 2750

GTTGCACCAG TTGGAGGAGC CCCGCACGTG GTTCGGCGTC CCGGGGGAGT 2800

ACCACTAGCA CCTCAAGACG TTCATGCCGT TGGAGAGGTT GAAGGACGCG 2850

CGGTTGCCCC TGCGGAAGTC GGGGACGCGC CTCTTCAGAG GGCTCGTCGC 2900

GCCTGCGAAG GCGCGGTACC ACCTCGAGCG GTCCGACCTA GCCTCCGCCG 2950

GGCCCTCGTC GCTGTCCCAG GAGAAGCGCG CCAAGAGCTT CTGGCTCCCG 3000

CCTCGCTCCG CCCGAAGAGG TCTGGTTCTT CGACTCCTGG ACACCGACTC 3050

GGGCGACTGG TACCTTCTAG AACAGACGAT GTCGAAGGTC CACCGGTCTC 3100

5 CCTACCTCAA GGACCGAAGG GCTTTCACGT AGGTGTCTCT GGACCGACGA 3150

GCCTTGTAAG ACGACAGCCT TTCGCTGCAC CACTTCTAGA CACTGAAACC 3200

GGAACGGGCC CTGTAGATGT TTCTGGGACT GATGCAGGCG TTCCCGTCAC 3250

GGGCCGACGG GGACTTCACC TACCGGGGAC TTTCGTAGAA GCTGTTCCAC 3300

ATGTGGTGCG TCTCACTGCA CACCAGGAAA CCCCACGAAG AGACCCTCTA 3350

10 GAAGAGAGAC CCCCAGAGGG GCATGGGACC CCACGTCTAG TTA CTCTCA 3400

AGACGGTCGC CGACTCTCTG CCGTGTTCCT ACTCCCGGGG CCTCGACCGG 3450

TGAGGGCGGT ATGCGGCGTA GTACGACTTG ACGACCAGGC CTCTGGGGTT 3500

CCGCTCTGGA CGTAAGAGCC TCGACCACCT CTAGGACCCC CTGGACGAGG 3550

TCCCGTCCCC GGACGTTCTC CTTCTCCTCC AGACGTACCG GGGCGCGTCG 3600

15 AGAGTCTCGA GTCTTCTCCC GTCGAAGAGC GTCCACAGGT GGTACCGGGA 3650

TGTGTAGCGG GTCCGACTGC GACTCCTGTC GGGCGGTTCG GACGTCGCGG 3700

TGTCGGACCG GCGGTCCATA ATGTTGACCC ACAGGAAAGG GCCCACGGAC 3750

CGGTCTCCC GACTCTGGGC ACCAAGGAGG TCCTACTTCT GTAAACTCCT 3800

TAAGGGGTAC TGGGGTTGCT GGATGTTTCC GAGACACCTG TTGGTCTGTC 3850

TGTCAACCTA CCACGACCGG AGCCTCCTCA AACTCGTCTA TCTCTCGTCC 3900
 GSTATCTGTTT TTTGCGCGAA GTCCATCGAC TTCGTCTCTC TCTCTTCCGT 3950
 CGTATGCAGT CGTAAAAGAA GAGACGTGAA TATTCTTTCT AGTTTCTGAA 4000
 ATTCTGAAAG CGATAAAGAA GACGATAGAT GATGTTTGAA GTTTCTCCTT 4050
 5 GGTCTCTCCG TTCTCCTCGT ACTTTCACCT GTTCCTCACA CTGGTGAATT 4100
 CGTGGTGTCC CTCCCCAATC CGGAGGCCTA CTGACGCCCC TCCGGACCTA 4150
 TTATAGGTCG GAGGGTGTTC TTCGACCACC TCGTCTCACA AGGGACTGAG 4200
 GAGGTTTCCTT TCCCTCTGCG GGAAAGTACC AGACGACTCA TTGTCCACGG 4250
 AAGGGTCTGT GACCGCAATG ACGAACTGGT TTCTCGGGAG TTCGCCGGGA 4300
 10 ATACGGTCGC ACTGTCTCCC GAGTGGAGAA CGGAAGATCC AGTGAAGAGT 4350
 GTTACAGGGA AGTCGTGGAC TGGGACACGG GCGGTCAATA AGGAACCATT 4400
 ATACTCATTG TGTAGTTTCT CATCA 4425

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 1298 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

20	Met	Gln	Arg	Gly	Ala	Ala	Leu	Cys	Leu	Arg	Leu	Trp	Leu	Cys	Leu	15
	1				5					10						
	Gly	Leu	Leu	Asp	Gly	Leu	Val	Ser	Gly	Tyr	Ser	Met	Thr	Pro	Pro	30
				20					25							
	Thr	Leu	Asn	Ile	Thr	Glu	Glu	Ser	His	Val	Ile	Asp	Thr	Gly	Asp	45
				35					40							
25	Ser	Leu	Ser	Ile	Ser	Cys	Arg	Gly	Gln	His	Pro	Leu	Glu	Trp	Ala	60
				50					55							

	Trp	Pro	Gly	Ala	Gln	Glu	Ala	Pro	Ala	Thr	Gly	Asp	Lys	Asp	Ser	
					65					70					75	
	Glu	Asp	Thr	Gly	Val	Val	Arg	Asp	Cys	Glu	Gly	Thr	Asp	Ala	Arg	
					80					85					90	
5	Pro	Tyr	Cys	Lys	Val	Leu	Leu	Leu	His	Glu	Val	His	Ala	Asn	Asp	
					95					100					105	
	Thr	Gly	Ser	Tyr	Val	Cys	Tyr	Tyr	Lys	Tyr	Ile	Lys	Ala	Arg	Ile	
					110					115					120	
10	Glu	Gly	Thr	Thr	Ala	Ala	Ser	Ser	Tyr	Val	Phe	Val	Arg	Asp	Phe	
					125					130					135	
	Glu	Gln	Pro	Phe	Ile	Asn	Lys	Pro	Asp	Thr	Leu	Leu	Val	Asn	Arg	
					140					145					150	
	Lys	Asp	Ala	Met	Trp	Val	Pro	Cys	Leu	Val	Ser	Ile	Pro	Gly	Leu	
					155					160					165	
15	Asn	Val	Thr	Leu	Arg	Ser	Gln	Ser	Ser	Val	Leu	Trp	Pro	Asp	Gly	
					170					175					180	
	Gln	Glu	Val	Val	Trp	Asp	Asp	Arg	Arg	Gly	Met	Leu	Val	Ser	Thr	
					185					190					195	
20	Pro	Leu	Leu	His	Asp	Ala	Leu	Tyr	Leu	Gln	Cys	Glu	Thr	Thr	Trp	
					200					205					210	
	Gly	Asp	Gln	Asp	Phe	Leu	Ser	Asn	Pro	Phe	Leu	Val	His	Ile	Thr	
					215					220					225	
	Gly	Asn	Glu	Leu	Tyr	Asp	Ile	Gln	Leu	Leu	Pro	Arg	Lys	Ser	Leu	
					230					235					240	
25	Glu	Leu	Leu	Val	Gly	Glu	Lys	Leu	Val	Leu	Asn	Cys	Thr	Val	Trp	
					245					250					255	
	Ala	Glu	Phe	Asn	Ser	Gly	Val	Thr	Phe	Asp	Trp	Asp	Tyr	Pro	Gly	
					260					265					270	
30	Lys	Gln	Ala	Glu	Arg	Gly	Lys	Trp	Val	Pro	Glu	Arg	Arg	Ser	Gln	
					275					280					285	
	Gln	Thr	His	Thr	Glu	Leu	Ser	Ser	Ile	Leu	Thr	Ile	His	Asn	Val	
					290					295					300	
	Ser	Gln	His	Asp	Leu	Gly	Ser	Tyr	Val	Cys	Lys	Ala	Asn	Asn	Gly	
					305					310					315	
35	Ile	Gln	Arg	Phe	Arg	Glu	Ser	Thr	Glu	Val	Ile	Val	His	Glu	Asn	
					320					325					330	
	Pro	Phe	Ile	Ser	Val	Glu	Trp	Leu	Lys	Gly	Pro	Ile	Leu	Glu	Ala	
					335					340					345	

	Thr Ala Gly Asp Glu Leu Val Lys Leu Pro Val Lys Leu Ala Ala	350	355	360
	Tyr Pro Pro Pro Glu Phe Gln Trp Tyr Lys Asp Gly Lys Ala Leu	365	370	375
5	Ser Gly Arg His Ser Pro His Ala Leu Val Leu Lys Glu Val Thr	380	385	390
	Glu Ala Ser Thr Gly Thr Tyr Thr Leu Ala Leu Trp Asn Ser Ala	395	400	405
10	Ala Gly Leu Arg Arg Asn Ile Ser Leu Glu Leu Val Val Asn Val	410	415	420
	Pro Pro Gln Ile His Glu Lys Glu Ala Ser Ser Pro Ser Ile Tyr	425	430	435
	Ser Arg His Ser Arg Gln Ala Leu Thr Cys Thr Ala Tyr Gly Val	440	445	450
15	Pro Leu Pro Leu Ser Ile Gln Trp His Trp Arg Pro Trp Thr Pro	455	460	465
	Cys Lys Met Phe Ala Gln Arg Ser Leu Arg Arg Arg Gln Gln Gln	470	475	480
20	Asp Leu Met Pro Gln Cys Arg Asp Trp Arg Ala Val Thr Thr Gln	485	490	495
	Asp Ala Val Asn Pro Ile Glu Ser Leu Asp Thr Trp Thr Glu Phe	500	505	510
	Val Glu Gly Lys Asn Lys Thr Val Ser Lys Leu Val Ile Gln Asn	515	520	525
25	Ala Asn Val Ser Ala Met Tyr Lys Cys Val Val Ser Asn Lys Val	530	535	540
	Gly Gln Asp Glu Arg Leu Ile Tyr Phe Tyr Val Thr Thr Ile Pro	545	550	555
30	Asp Gly Phe Thr Ile Glu Ser Lys Pro Ser Glu Glu Leu Leu Glu	560	565	570
	Gly Gln Pro Val Leu Leu Ser Cys Gln Ala Asp Ser Tyr Lys Tyr	575	580	585
	Glu His Leu Arg Trp Tyr Arg Leu Asn Leu Ser Thr Leu His Asp	590	595	600
35	Ala His Gly Asn Pro Leu Leu Leu Asp Cys Lys Asn Val His Leu	605	610	615
	Phe Ala Thr Pro Leu Ala Ala Ser Leu Glu Glu Val Ala Pro Gly	620	625	630

	Ala Arg His Ala Thr Leu Ser Leu Ser Ile Pro Arg Val Ala Pro	635	640	645
	Glu His Glu Gly His Tyr Val Cys Glu Val Gln Asp Arg Arg Ser	650	655	660
5	His Asp Lys His Cys His Lys Lys Tyr Leu Ser Val Gln Ala Leu	665	670	675
	Glu Ala Pro Arg Leu Thr Gln Asn Leu Thr Asp Leu Leu Val Asn	680	685	690
10	Val Ser Asp Ser Leu Glu Met Gln Cys Leu Val Ala Gly Ala His	695	700	705
	Ala Pro Ser Ile Val Trp Tyr Lys Asp Glu Arg Leu Leu Glu Glu	710	715	720
	Lys Ser Gly Val Asp Leu Ala Asp Ser Asn Gln Lys Leu Ser Ile	725	730	735
15	Gln Arg Val Arg Glu Glu Asp Ala Gly Arg Tyr Leu Cys Ser Val	740	745	750
	Cys Asn Ala Lys Gly Cys Val Asn Ser Ser Ala Ser Val Ala Val	755	760	765
20	Glu Gly Ser Glu Asp Lys Gly Ser Met Glu Ile Val Ile Leu Val	770	775	780
	Gly Thr Gly Val Ile Ala Val Phe Phe Trp Val Leu Leu Leu Leu	785	790	795
	Ile Phe Cys Asn Met Arg Arg Pro Ala His Ala Asp Ile Lys Thr	800	805	810
25	Gly Tyr Leu Ser Ile Ile Met Asp Pro Gly Glu Val Pro Leu Glu	815	820	825
	Glu Gln Cys Glu Tyr Leu Ser Tyr Asp Ala Ser Gln Trp Glu Phe	830	835	840
30	Pro Arg Glu Arg Leu His Leu Gly Arg Val Leu Gly Tyr Gly Ala	845	850	855
	Phe Gly Lys Val Val Glu Ala Ser Ala Phe Gly Ile His Lys Gly	860	865	870
	Ser Ser Cys Asp Thr Val Ala Val Lys Met Leu Lys Glu Gly Ala	875	880	885
35	Thr Ala Ser Glu His Arg Ala Leu Met Ser Glu Leu Lys Ile Leu	890	895	900
	Ile His Ile Gly Asn His Leu Asn Val Val Asn Leu Leu Gly Ala	905	910	915

	Cys Thr Lys Pro Gln Gly Pro Leu Met Val Ile Val Glu Phe Cys	920	925	930
	Lys Tyr Gly Asn Leu Ser Asn Phe Leu Arg Ala Lys Arg Asp Ala	935	940	945
5	Phe Ser Pro Cys Ala Glu Lys Ser Pro Glu Gln Arg Gly Arg Phe	950	955	960
	Arg Ala Met Val Glu Leu Ala Arg Leu Asp Arg Arg Arg Pro Gly	965	970	975
10	Ser Ser Asp Arg Val Leu Phe Ala Arg Phe Ser Lys Thr Glu Gly	980	985	990
	Gly Ala Arg Arg Ala Ser Pro Asp Gln Glu Ala Glu Asp Leu Trp	995	1000	1005
	Leu Ser Pro Leu Thr Met Glu Asp Leu Val Cys Tyr Ser Phe Gln	1010	1015	1020
15	Val Ala Arg Gly Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His	1025	1030	1035
	Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Ser Asp Val	1040	1045	1050
20	Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp	1055	1060	1065
	Pro Asp Tyr Val Arg Lys Gly Ser Ala Arg Leu Pro Leu Lys Trp	1070	1075	1080
	Met Ala Pro Glu Ser Ile Phe Asp Lys Val Tyr Thr Thr Gln Ser	1085	1090	1095
25	Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu	1100	1105	1110
	Gly Ala Ser Pro Tyr Pro Gly Val Gln Ile Asn Glu Glu Phe Cys	1115	1120	1125
30	Gln Arg Leu Arg Asp Gly Thr Arg Met Arg Ala Pro Glu Leu Ala	1130	1135	1140
	Thr Pro Ala Ile Arg Arg Ile Met Leu Asn Cys Trp Ser Gly Asp	1145	1150	1155
	Pro Lys Ala Arg Pro Ala Phe Ser Glu Leu Val Glu Ile Leu Gly	1160	1165	1170
35	Asp Leu Leu Gln Gly Arg Gly Leu Gln Glu Glu Glu Val Cys	1175	1180	1185
	Met Ala Pro Arg Ser Ser Gln Ser Ser Glu Glu Gly Ser Phe Ser	1190	1195	1200

Gln Val Ser Thr Met Ala Leu His Ile Ala Gln Ala Asp Ala Glu
 1205 1210 1215

Asp Ser Pro Pro Ser Leu Gln Arg His Ser Leu Ala Ala Arg Tyr
 1220 1225 1230

5 Tyr Asn Trp Val Ser Phe Pro Gly Cys Leu Ala Arg Gly Ala Glu
 1235 1240 1245

Thr Arg Gly Ser Ser Arg Met Lys Thr Phe Glu Glu Phe Pro Met
 1250 1255 1260

10 Thr Pro Thr Thr Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp Ser
 1265 1270 1275

Gly Met Val Leu Ala Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg
 1280 1285 1290

His Arg Gln Glu Ser Gly Phe Arg
 1295 1298

15 (2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3348 bases
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ATGGCTGGGA TTTTCTATTT CGCCCTATTT TCGTGTCTCT TCGGGATTTG 50

CGACGCTGTC ACAGGTTCCA GGGTATACCC CGCGAATGAA GTTACCTTAT 100

TGGATTCCAG ATCTGTTCAG GGAGAACTTG GGTGGATAGC AAGCCCTCTG 150

25 GAAGGAGGGT GGGAGGAAGT GAGTATCATG GATGAAAAAA ATACACCAAT 200

CCGAACCTAC CAAGTGTGCA ATGTGATGGA ACCCAGCCAG AATAACTGGC 250

TACGAAGTGA TTGGATCACC CGAGAAGGGG CTCAGAGGGT GTATATTGAG 300

ATTAAATTCA CCTTGAGGGA CTGCAATAGT CTTCCGGGCG TCATGGGGAC 350

TTGCAAGGAG ACGTTTAACC TGTACTACTA TGAATCAGAC AACGACAAAG 400

30 AGCGTTTCAT CAGAGAGAAC CAGTTTGTCA AAATTGACAC CATTGCTGCT 450